33358, A NOVEL HUMAN ANKYRIN FAMILY MEMBER AND USES THEREOF

Related Applications

This application claims priority to U.S. Provisional Patent Application No. 60/212,222 filed on June 16, 2000, incorporated herein in its entirety by reference.

Background of the Invention

Protein-protein interactions are critical for virtually all cellular processes. Cell growth, differentiation, and death are mechanisms regulated by the interaction of proteins with one another. Proteins altered in binding specificity may lead to aberrant or absent interactions and are responsible for a variety of diseases; *e.g.* birth defects, cancer, and heart disease. Various motifs mediating such interactions have been identified in recent years and include death domains, PDZ domains, WW domains, leucine zippers and leucine rich repeats, and ankyrin repeats.

Ankyrin repeat containing proteins are a diverse family of proteins which include cell cycle proteins, transcription factors, and proteins that mediate development (Blank, V. *et al.* (1992) *Trends Biochem. Sci.* 17:135-140, Bork, P. (1993) *Proteins* 17:363-374). Ankyrin repeats are named for their homology to repeats in the erythrocyte protein ankyrin. Such repeats are 33 amino acids long and are typically found in clusters of four or more. The structure of ankyrin-repeat regions of many proteins have been solved and it is well documented that each ankyrin-repeat forms an L shaped structure whereby two alpha-helices are connected by a beta-hairpin (a helix-loop-helix) (Batchelor, A.H. *et al.* (1998) *Science* 279:1037-1041, Zhang Z. *et al.* (1998) *J. Biol. Chem.* 273:18681-18684, Jacobs M.D. *et al.* (1998) *Cell* 95:749-758). The alpha helices are often stacked upon one another forming a scaffold by which the beta-hairpin is exposed and available to bind heterologous proteins.

Ankyrin-repeat containing proteins are present in nearly all cells. These proteins have been identified as important for diverse activities including regulation of cardiac cellular processes; *e.g.* cardiogenesis and heart diseases (Zou, Y. *et al.* (1997) *Development* 124:793-804, Yang, Y. *et al.* (1998) *Structure* 15:619-626, Kuo, H. *et al.* (1999) *Development* 126:4223-4234).

Summary of the Invention

The present invention is based, at least in part, on the discovery of ankyrin repeat-containing protein family members, referred to herein as "Cardiac/Skeletal Muscle-Restricted Ankyrin-Repeat Containing Protein" or "C/SKARP" nucleic acid and protein molecules. The C/SKARP nucleic acid and protein molecules of the present invention are

useful as modulating agents in regulating a variety of cellular processes, *e.g.*, myogenic cellular processes including, but not limited to cardiac cellular processes. Accordingly, in one aspect, this invention provides isolated nucleic acid molecules encoding C/SKARP proteins or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection of C/SKARP-encoding nucleic acids.

In one embodiment, the invention features an isolated nucleic acid molecule that includes the nucleotide sequence set forth in SEQ ID NO:1 or SEQ ID NO:3. In another embodiment, the invention features an isolated nucleic acid molecule that encodes a polypeptide including the amino acid sequence set forth in SEQ ID NO:2. In another embodiment, the invention features an isolated nucleic acid molecule that includes the nucleotide sequence contained in the plasmid deposited with ATCC® as Accession Number _____.

In still other embodiments, the invention features isolated nucleic acid molecules including nucleotide sequences that are substantially identical (*e.g.*, 60% identical) to the nucleotide sequence set forth as SEQ ID NO:1 or SEQ ID NO:3. The invention further features isolated nucleic acid molecules including at least 30 contiguous nucleotides of the nucleotide sequence set forth as SEQ ID NO:1 or SEQ ID NO:3. In another embodiment, the invention features isolated nucleic acid molecules which encode a polypeptide including an amino acid sequence that is substantially identical (*e.g.*, 60% identical) to the amino acid sequence set forth as SEQ ID NO:2. Also featured are nucleic acid molecules which encode allelic variants of the polypeptide having the amino acid sequence set forth as SEQ ID NO:2. In addition to isolated nucleic acid molecules encoding full-length polypeptides, the present invention also features nucleic acid molecules which encode fragments, for example biologically active or antigenic fragments, of the full-length polypeptides of the present invention (*e.g.*, fragments including at least 10 contiguous amino acid residues of the amino acid sequence of SEQ ID NO:2). In still other embodiments, the invention features isolated nucleic acid molecules that are complementary to, are antisense to, or hybridize under stringent conditions to the isolated nucleic acid molecules described herein.

In a related aspect, the invention provides vectors including the isolated nucleic acid molecules described herein (*e.g.*, C/SKARP-1-encoding nucleic acid molecules). Such vectors can optionally include nucleotide sequences encoding heterologous polypeptides. Also featured are host cells including such vectors (*e.g.*, host cells including vectors suitable for producing C/SKARP-1 nucleic acid molecules and polypeptides).

In another aspect, the invention features isolated C/SKARP-1 polypeptides and/or biologically active or antigenic fragments thereof. Exemplary embodiments feature a polypeptide including the amino acid sequence set forth as SEQ ID NO:2, a polypeptide including an amino acid sequence at least 60% identical to the amino acid sequence set forth as SEQ ID NO:2, a polypeptide encoded by a nucleic acid molecule including a nucleotide

sequence at least 60% identical to the nucleotide sequence set forth as SEQ ID NO:1 or SEQ ID NO:3. Also featured are fragments of the full-length polypeptides described herein (*e.g.*, fragments including at least 10 contiguous amino acid residues of the sequence set forth as SEQ ID NO:2) as well as fragments of allelic variants of the polypeptide having the amino acid sequence set forth as SEQ ID NO:2.

The C/SKARP-1 polypeptides and/or biologically active or antigenic fragments thereof, are useful, for example, as reagents or targets in assays applicable to treatment and/or diagnosis of C/SKARP-1 mediated or related disorders. In one embodiment, a C/SKARP-1 polypeptide or fragment thereof has a C/SKARP-1 activity. In another embodiment, a C/SKARP-1 polypeptide or fragment thereof has an ankyrin repeat domain and optionally, has a C/SKARP-1 activity. In a related aspect, the invention features antibodies (*e.g.*, antibodies which specifically bind to any one of the polypeptides, as described herein) as well as fusion polypeptides including all or a fragment of a polypeptide described herein.

The present invention further features methods for detecting C/SKARP-1 polypeptides and/or C/SKARP-1 nucleic acid molecules, such methods featuring, for example, a probe, primer or antibody described herein. Also featured are kits for the detection of C/SKARP-1 polypeptides and/or C/SKARP-1 nucleic acid molecules. In a related aspect, the invention features methods for identifying compounds which bind to and/or modulate the activity of a C/SKARP-1 polypeptide or C/SKARP-1 nucleic acid molecule described herein. Further featured are methods for modulating a C/SKARP-1 activity.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

Brief Description of the Drawings

Figure 1A-1B depicts a cDNA sequence (SEQ ID NO:1) and predicted amino acid sequence (SEQ ID NO:2) of human C/SKARP-1. The methionine-initiated open reading frame of human C/SKARP-1 (without the 5' and 3' untranslated regions) starts at nucleotide 75 until the termination codon (shown also as coding sequence SEQ ID NO:3).

Figure 2A-2B depicts C/SKARP-1 mRNA expression by probing a library array using RT-PCR.

Figure 3 depicts a structural, hydrophobicity, and antigenicity analysis of the human C/SKARP-1 protein.

Detailed Description of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as "Cardiac/Skeletal Muscle Restricted Ankyrin-Repeat Containing Protein" or "C/SKARP" protein and nucleic acid molecules, which comprise a family of molecules having certain conserved structural and functional features.

The term "family" when referring to the protein and nucleic acid molecules of the invention is intended to mean two or more proteins or nucleic acid molecules having at least one common structural domain or motif and having sufficient amino acid or nucleotide sequence homology or identity as defined herein. Such family members can be naturally or non-naturally occurring and can be from either the same or different species. For example, a family can contain a first protein of human origin, as well as other, distinct proteins of human origin or alternatively, can contain homologues of non-human origin. Members of a family may also have common functional characteristics.

For example, a C/SKARP protein of the present invention can include at least one "ankyrin repeat domain" in the polypeptide (or encoded by the corresponding nucleic acid sequence). As used herein, the term "ankyrin repeat domain" includes a protein domain involved in protein-protein interactions having an amino acid sequence of about 190 to 200 (*e.g.*, about 196) amino acid residues in length and including six ankyrin repeats (*e.g.*, including six consecutive copies of an ankyrin repeat). In another embodiment, an ankyrin repeat domain includes at least about 160 to about 170 (*e.g.*, about 163 to 164) amino acid residues, about 125 to 135 (*e.g.*, about 130 to 131 amino acid residues, about 90 to 100 (*e.g.*, about 95 to 99) amino acid residues or about 60 to 70 (*e.g.*, about 65 to 66) amino acid residues and includes five, four, three or two ankyrin repeats, respectively.

In a preferred embodiment, a C/SKARP polypeptide or protein has an "ankyrin repeat domain" which includes at least about 190 to 200, about 160 to 170, or about 125 to 135 amino acid residues and has at least about 60%, 70% 80% 90% 95%, 99%, or 100% identity with the "ankyrin repeat domain," of human C/SKARP-1 (*e.g.*, amino acids 64 to 259 of SEQ ID NO:2).

As used herein, the term "ankyrin repeat" includes a protein motif typically containing about 33 amino acid residues, initially identified in ankyrin and now identified in over 650 distinct proteins and known to have a role in protein-protein interactions (see *e.g.*, Bork (1993) *Proteins: Structure, Function, and Genetics* 17:363-374). Preferably, an ankyrin repeat has an amino acid sequence of about 25-40 amino acid residues and has a bit score for the alignment of the sequence to an ankyrin repeat (HMM) (*e.g.*, the Pfam ankyrin repeat HMM having Accession Number PF00023) of at least 10. More preferably, an ankyrin repeat includes at least about 30-36, about 31-35 amino acid residues, about 32-34, or typically about 33 amino acid residues, and has a bit score for the alignment of the sequence to an ankyrin repeat (HMM) of at least 12, 14, 16, 18, 20, 22, 24, 26, or greater. In

a preferred embodiment, a C/SKARP protein of the present invention has at least one, and preferably two, three, four, five, or most preferably, six or more ankyrin repeats, as defined herein.

To identify the presence of an ankyrin repeat in a C/SKARP-1 protein, and make the determination that a query protein has a particular profile, the amino acid sequence of the protein is searched against a database of HMMs (*e.g.*, the Pfam database, release 5.3) using the default parameters (http://www.sanger.ac.uk/Software/Pfam/HMM_search). For example, the search can be performed using the hmmsf program (family specific) using the default parameters (*e.g.*, a threshold score of 15) for determining a hit. hmmsf is available as part of the HMMER package of search programs (HMMER 2.1.1, Dec. 1998) which is freely distributed by the Washington University School of Medicine. Alternatively, the threshold score for determining a hit can be lowered (*e.g.*, to 8 bits). A description of the Pfam database can be found in Sonhammer *et al.* (1997) *Proteins* 28(3)405-420 and a detailed description of HMMs can be found, for example, in Gribskov *et al.*(1990) *Meth. Enzymol.* 183:146-159; Gribskov *et al.*(1987) *Proc. Natl. Acad. Sci. USA* 84:4355-4358; Krogh *et al.*(1994) *J. Mol. Biol.* 235:1501-1531; and Stultz *et al.*(1993) *Protein Sci.* 2:305-314, the contents of which are incorporated herein by reference.

A search was performed against the HMM database resulting in the identification of six ankyrin repeats in the amino acid sequence of human C/SKARP-1 (SEQ ID NO:2) at about residues 64-96, 97-129, 130-162, 165-194, 195-227, and 229-259 of SEQ ID NO:2. Identification of ankyrin repeats in a C/SKARP protein of the present invention according to the above-described methodologies further facilitates identification of an ankyrin repeat domain, *e.g.*, comprising six ankyrin repeats as defined herein.

In yet another embodiment, C/SKARP-1 family members include at least one or more transmembrane domains. As used herein, a "transmembrane domain" includes a protein domain having at least about 10 amino acid residues of which about 60% of the amino acid residues contain non-polar side chains, for example, alanine, valine, leucine, isoleucine, proline, phenylalanine, tryptophan, and methionine. In a preferred embodiment, a "transmembrane domain" includes a protein domain having at least about 13, preferably about 16, more preferably about 19, and even more preferably about 21, 23, 25, 30, 35 or 40 amino acid residues, of which at least about 70%, preferably about 80%, and more preferably about 90% of the amino acid residues contain non-polar side chains, for example, alanine, valine, leucine, isoleucine, proline, phenylalanine, tryptophan, and methionine. A transmembrane domain is lipophillic in nature. Predicted transmembrane domains are found, for example, from about amino acid residues 13-35 and 135-151 of SEQ ID NO:2

In yet another embodiment, C/SKARP-1 family members includes a signal peptide. As used herein, a "signal sequence" includes a peptide of at least about 20 amino acid residues in length which occurs at the N-terminus of secretory and integral membrane

proteins and which contains at least 55% hydrophobic amino acid residues. In a preferred embodiment, a signal sequence contains at least about 15-45 amino acid residues, preferably about 20-42 amino acid residues. Signal sequences of 25-35 amino acid residues and 28-32 amino acid residues are also within the scope of the invention. As used herein, a signal sequence has at least about 40-70%, preferably about 50-65%, and more preferably about 55-60% hydrophobic amino acid residues (*e.g.*, Alanine, Valine, Leucine, Isoleucine, Phenylalanine, Tyrosine, Tryptophan, or Proline). Such a "signal sequence", also referred to in the art as a "signal peptide", serves to direct a protein containing such a sequence to a lipid bilayer. A predicted signal peptide is found, for example, from about amino acid residues 1-43 of SEQ ID NO:2 (although this possible signal peptide is not believed to be utilized by the C/SKARP-1 polypeptide of SEQ ID NO:2).

Isolated proteins of the present invention, for example C/SKARP proteins, preferably have an amino acid sequence sufficiently identical to the amino acid sequence of SEQ ID NO:2, or are encoded by a nucleotide sequence sufficiently identical to SEQ ID NO:1 or 3. As used herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (*e.g.*, an amino acid residue which has a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences share common structural domains or motifs and/or a common functional activity. For example, amino acid or nucleotide sequences which share common structural domains having at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homology or identity across the amino acid sequences of the domains and contain at least one and preferably two structural domains or motifs, are defined herein as sufficiently identical. Furthermore, amino acid or nucleotide sequences which share at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homology or identity and share a common functional activity are defined herein as sufficiently identical.

In a preferred embodiment, a C/SKARP protein includes at least one or more ankyrin repeat domain, and has an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more homologous or identical to the amino acid sequence of SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In yet another preferred embodiment, a C/SKARP protein includes at least one or more ankyrin repeat domain, and is encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3. In another preferred embodiment, a C/SKARP protein includes at least one or more ankyrin repeat domain, and has a C/SKARP activity.

As used herein, a "C/SKARP activity", "biological activity of C/SKARP" or "functional activity of C/SKARP", refers to an activity exerted by a C/SKARP protein, polypeptide or nucleic acid molecule on *e.g.*, a C/SKARP-responsive cell or on a C/SKARP target, *e.g.*, a protein activity, as determined *in vivo* or *in vitro*. In one embodiment, a C/SKARP activity is a direct activity, such as an association with a C/SKARP target molecule. A "target molecule" or "binding partner" is a molecule with which a C/SKARP protein binds or interacts in nature. In an exemplary embodiment, a C/SKARP target molecule is a protein molecule (*e.g.*, a second C/SKARP protein or a non-C/SKARP protein molecule). A C/SKARP activity can also be an indirect activity, *e.g.*, a cellular signaling activity mediated by interaction of the C/SKARP protein with a C/SKARP target. In a preferred embodiment, the C/SKARP proteins of the present invention have one or more of the following activities: (i) mediation of specific macromoleculer interactions; (ii) mediation of interactions between proteins and/or between regions of a single protein; (iii) formation of binding sites for distinct proteins (*e.g.*, non-C/SKARP proteins); (iv) bridging of cellular components; (v) regulation of gene expression (*e.g.*, cardiac gene expression); (vi) modulation of cellular localization (*e.g.*, anchoring C/SKARP binding proteins in a specific cellular localization); (vii) modulation of development and/or differentiation (*e.g.*, myogenic development and/or differentiation, heart development and/or differentiation); (viii) modulation of cardiac maturation and/or morphogenesis; (ix) as a marker (*e.g.*, an early marker) of cardiac and/or myogenic cell lineage; and (x) modulation and/or treatment of cardiac hypertrophy.

Inhibition or over stimulation of the activity of proteins involved in signaling pathways associated with cellular growth can lead to perturbed cellular growth, which can in turn lead to cellular growth related disorders. As used herein, a "cellular growth related disorder" includes a disorder, disease, or condition characterized by a deregulation, *e.g.*, an upregulation or a downregulation, of cellular growth. Cellular growth deregulation may be due to a deregulation of cellular proliferation, cell cycle progression, cellular differentiation and/or cellular hypertrophy. Examples of cellular growth related disorders include cardiovascular disorders such as heart failure, hypertension, atrial fibrillation, dilated cardiomyopathy, idiopathic cardiomyopathy, or angina; proliferative disorders or differentiative disorders such as cancer, *e.g.*, melanoma, prostate cancer, cervical cancer, breast cancer, colon cancer, or sarcoma.

As used herein, the term "cardiovascular disorder" includes a disease, disorder, or state involving the cardiovascular system, *e.g.*, the heart, the blood vessels, and/or the blood. A cardiovascular disorder can be caused by an imbalance in arterial pressure, a malfunction of the heart, or an occlusion of a blood vessel, *e.g.*, by a thrombus. Examples of such disorders include hypertension, atherosclerosis, coronary artery spasm, coronary artery disease, valvular disease, arrhythmias, and cardiomyopathies.

As used herein, the term "congestive heart failure" includes a condition characterized by a diminished capacity of the heart to supply the oxygen demands of the body. Symptoms and signs of congestive heart failure include diminished blood flow to the various tissues of the body, accumulation of excess blood in the various organs, *e.g.*, when the heart is unable to pump out the blood returned to it by the great veins, exertional dyspnea, fatigue, and/or peripheral edema, *e.g.*, peripheral edema resulting from left ventricular dysfunction. Congestive heart failure may be acute or chronic. The manifestation of congestive heart failure usually occurs secondary to a variety of cardiac or systemic disorders that share a temporal or permanent loss of cardiac function. Examples of such disorders include hypertension, coronary artery disease, valvular disease, and cardiomyopathies, *e.g.*, hypertrophic, dilative, or restrictive cardiomyopathies. Congestive heart failure is described in, for example, Cohn J.N. *et al.* (1998) *American Family Physician* 57:1901-04, the contents of which are incorporated herein by reference.

A partial human C/SKARP -1 cDNA has been identified, which is approximately 1538 nucleotides in length, encodes a protein which is approximately 323 amino acid residues in length.

A plasmid containing the nucleotide sequence encoding human C/SKARP-1 was deposited with American Type Culture Collection (ATCC), 10801 University Boulevard, Manassas, VA 20110-2209, on _____ and assigned Accession Number _____. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

Various aspects of the invention are described in further detail in the following subsections:

I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode C/SKARP-1 proteins or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes to identify C/SKARP-1-encoding nucleic acid molecules (*e.g.*, C/SKARP-1 mRNA) and fragments for use as PCR primers for the amplification or mutation of C/SKARP-1 nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (*e.g.*, cDNA or genomic DNA) and RNA molecules (*e.g.*, mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

The term "isolated nucleic acid molecule" includes nucleic acid molecules which are separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. For example, with regards to genomic DNA, the term "isolated" includes nucleic acid molecules which are separated from the chromosome with which the genomic DNA is naturally associated. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (*i.e.*, sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated C/SKARP-1 nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, *e.g.*, a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1 or 3,or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or portion of the nucleic acid sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, as a hybridization probe, C/SKARP-1 nucleic acid molecules can be isolated using standard hybridization and cloning techniques (*e.g.*, as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ can be isolated by the polymerase chain reaction (PCR) using synthetic oligonucleotide primers designed based upon the sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

A nucleic acid of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to C/SKARP-1 nucleotide sequences can be prepared by standard synthetic techniques, *e.g.*, using an automated DNA synthesizer.

In a one embodiment, an isolated nucleic acid molecule of the invention comprises the nucleotide sequence shown in SEQ ID NO:1. The sequence of SEQ ID NO:1 corresponds to the human C/SKARP-1 cDNA. This cDNA comprises sequences encoding the human C/SKARP-1 protein (*i.e.*, "the coding region", from nucleotides 75-1046), as well as 5' untranslated sequences (nucleotides 1-74) and 3' untranslated sequences (nucleotides 1047-1538). Alternatively, the nucleic acid molecule can comprise only the coding region of SEQ ID NO:1 (*e.g.*, nucleotides 75-1046, corresponding to SEQ ID NO:3). Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention comprises SEQ ID NO:3 and nucleotides 1-74 of SEQ ID NO:1. In yet another embodiment, the isolated nucleic acid molecule comprises SEQ ID NO:3 and nucleotides 1047-1538 of SEQ ID NO:1. In yet another embodiment, the nucleic acid molecule consists of the nucleotide sequence set forth as SEQ ID NO:1 or SEQ ID NO:3.

In still another embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ , or a portion of any of these nucleotide sequences. A nucleic acid molecule which is complementary to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, is one which is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more identical to the nucleotide sequence shown in SEQ ID NO:1 or 3 (*e.g.*, to the entire length of the nucleotide sequence), or to the nucleotide sequence (*e.g.*, the entire length of the nucleotide sequence) of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or to a portion or complement of any of these nucleotide sequences. In one embodiment, a nucleic acid molecule of the present invention comprises a nucleotide sequence which is at least (or no greater than) 50-100, 100-250, 250-500, 500-750, 750-1000, 1000-1250, 1250-1500, 1500-1700 or more nucleotides in length and hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the nucleic acid sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, for example, a fragment which can be used as a probe or primer or a fragment encoding a portion of a C/SKARP-1 protein, *e.g.*, a biologically active portion of a C/SKARP-1 protein. The nucleotide sequence determined from the cloning of the C/SKARP-1 gene allows for the generation of probes and primers designed for use in identifying and/or cloning other C/SKARP-1 family members, as well as C/SKARP-1 homologues from other species. The probe/primer (*e.g.*, oligonucleotide) typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 consecutive nucleotides of a sense sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, of an anti-sense sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____.

Exemplary probes or primers are at least (or no greater than)12 or 15, 20 or 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75 or more nucleotides in length and/or comprise consecutive nucleotides of an isolated nucleic acid molecule described herein. Also included within the scope of the present invention are probes or primers comprising contiguous or consecutive nucleoitdes of an isolated nucleic acid molecule described herein, but for the difference of 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 bases within the probe or primer sequence. Probes based on the C/SKARP-1 nucleotide sequences can be used to detect (*e.g.*, specifically detect) transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.*, the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. In another embodiment a set of primers is provided, *e.g.*, primers suitable for use in a PCR, which can be used to amplify a selected region of a C/SKARP-1 sequence, *e.g.*, a domain, region, site or other sequence described herein. The primers should be at least 5, 10, or 50 base pairs in length and less than 100, or less than 200, base pairs in length. The primers should be identical, or differs by no greater than 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 bases when compared to a sequence disclosed herein or to the sequence of a naturally occurring variant. Such probes can be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a C/SKARP-1 protein, such as by measuring a level of a C/SKARP-1-encoding nucleic acid in a sample of cells from a subject *e.g.*, detecting C/SKARP-1 mRNA levels or determining whether a genomic C/SKARP-1 gene has been mutated or deleted.

A nucleic acid fragment encoding a "biologically active portion of a C/SKARP-1 protein" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, which encodes a polypeptide having a C/SKARP-1 biological activity (the biological activities of the C/SKARP-1 proteins are described herein), expressing the encoded portion of the C/SKARP-1 protein (*e.g.*, by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the C/SKARP-1 protein. In an exemplary embodiment, the nucleic acid molecule is at least 50-100, 100-250, 250-500, 500-700, 750-1000, 1000-1250, 1250-1500, 1500-1700 or more nucleotides in length and encodes a protein having a GPCR52871 activity (as described herein).

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, due to degeneracy of the genetic code and thus encode the same C/SKARP-1 proteins as those encoded by the nucleotide sequence shown in SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence which differs by at least 1, but no greater than 5, 10, 20, 50 or 100 amino acid residues from the amino acid sequence shown in SEQ ID NO:2, or the amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number _____. In yet another embodiment, the nucleic acid molecule encodes the amino acid sequence of human GPCR52871. If an alignment is needed for this comparison, the sequences should be aligned for maximum homology.

Nucleic acid variants can be naturally occurring, such as allelic variants (same locus), homologues (different locus), and orthologues (different organism) or can be non naturally occurring. Non-naturally occurring variants can be made by mutagenesis techniques, including those applied to polynucleotides, cells, or organisms. The variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared in the encoded product).

Allelic variants result, for example, from DNA sequence polymorphisms within a population (*e.g.*, the human population) that lead to changes in the amino acid sequences of the C/SKARP-1 proteins. Such genetic polymorphism in the C/SKARP-1 genes may exist among individuals within a population due to natural allelic variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules which include an open reading frame encoding a C/SKARP-1 protein, preferably a mammalian C/SKARP-1 protein, and can further include non-coding regulatory sequences, and introns.

Accordingly, in one embodiment, the invention features isolated nucleic acid molecules which encode a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with ATCC as Accession Number _____, wherein the nucleic acid molecule hybridizes to a complement of a nucleic acid molecule comprising SEQ ID NO:1 or SEQ ID NO:3, for example, under stringent hybridization conditions.

Allelic variants of human C/SKARP-1 include both functional and non-functional C/SKARP-1 proteins. Functional allelic variants are naturally occurring amino acid sequence variants of the human C/SKARP-1 protein that maintain the ability to bind a C/SKARP-1 ligand and/or modulate cellular mechanisms associated with cell growth or differentiation. Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2 or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein.

Non-functional allelic variants are naturally occurring amino acid sequence variants of the human C/SKARP-1 protein that do not have the ability to either bind a C/SKARP-1 ligand and/or modulate cellular mechanisms associated with cell growth or differentiation. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion or premature truncation of the amino acid sequence of SEQ ID NO:2 or a substitution, insertion or deletion in critical residues or critical regions.

The present invention further provides non-human orthologues (*e.g.*, non-human orthologues of the human C/SKARP-1 protein). Orthologues of the human C/SKARP-1 protein are proteins that are isolated from non-human organisms and possess the same C/SKARP-1 ligand binding and/or modulation of cellular mechanisms associated with cell growth or differentiation of the human C/SKARP-1 protein. Orthologues of the human C/SKARP-1 protein can readily be identified as comprising an amino acid sequence that is substantially homologous to SEQ ID NO:2.

Moreover, nucleic acid molecules encoding other C/SKARP-1 family members and, thus, which have a nucleotide sequence which differs from the C/SKARP-1 sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, another C/SKARP-1 cDNA can be identified based on the nucleotide sequence of human C/SKARP-1. Moreover, nucleic acid molecules encoding C/SKARP-1 proteins from different species, and which, thus, have a nucleotide sequence which differs from the C/SKARP-1 sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ are intended to be within the scope of the invention. For example, a mouse C/SKARP-1 cDNA can be identified based on the nucleotide sequence of a human C/SKARP-1.

Nucleic acid molecules corresponding to natural allelic variants and homologues of the C/SKARP-1 cDNAs of the invention can be isolated based on their homology to the C/SKARP-1 nucleic acids disclosed herein using the cDNAs disclosed herein, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Nucleic acid molecules corresponding to natural allelic variants and homologues of the C/SKARP-1 cDNAs of the invention can further be isolated by mapping to the same chromosome or locus as the C/SKARP-1 gene.

Orthologues, homologues and allelic variants can be identified using methods known in the art (*e.g.*, by hybridization to an isolated nucleic acid molecule of the present invention, for example, under stringent hybridization conditions). In one embodiment, an isolated nucleic acid molecule of the invention is at least 15, 20, 25, 30 or more nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. In another embodiment, the nucleic acid is at least 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 467, 500, 550, 600, 650, 700, 750, 800, 850, 900, or 950 nucleotides in length.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences that are significantly identical or homologous to each other remain hybridized to each other. Preferably, the conditions are such that sequences at least about 70%, more preferably at least about 80%, even more preferably at least about 85% or 90% identical to each other remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, Ausubel *et al.*, eds., John Wiley & Sons, Inc. (1995), sections 2, 4 and 6. Additional stringent conditions can be found in *Molecular Cloning: A Laboratory Manual*, Sambrook *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989), chapters 7, 9 and 11. A preferred, non-limiting example of stringent hybridization conditions includes hybridization in 4X sodium chloride/sodium citrate (SSC), at about 65-70°C (or alternatively hybridization in 4X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 1X SSC, at about 65-70°C. A preferred, non-limiting example of highly stringent hybridization conditions includes hybridization in 1X SSC, at about 65-70°C (or alternatively hybridization in 1X SSC plus 50% formamide at about 42-50°C) followed by one or more washes in 0.3X SSC, at about 65-70°C. A preferred, non-limiting example of reduced stringency hybridization conditions includes hybridization in 4X SSC, at about 50-60°C (or alternatively hybridization in 6X SSC plus 50% formamide at about 40-45°C) followed by one or more washes in 2X SSC, at about 50-60°C. Ranges intermediate to the above-recited values, *e.g.*, at 65-70°C or at 42-50°C are also intended to be encompassed by the present invention. SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4)

can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes each after hybridization is complete. The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, $T_m(°C) = 2(\#\text{ of A + T bases}) + 4(\#\text{ of G + C bases})$. For hybrids between 18 and 49 base pairs in length, $T_m(°C) = 81.5 + 16.6(\log_{10}[Na^+]) + 0.41(\%G+C) - (600/N)$, where N is the number of bases in the hybrid, and $[Na^+]$ is the concentration of sodium ions in the hybridization buffer ($[Na^+]$ for 1xSSC = 0.165 M). It will also be recognized by the skilled practitioner that additional reagents may be added to hybridization and/or wash buffers to decrease non-specific hybridization of nucleic acid molecules to membranes, for example, nitrocellulose or nylon membranes, including but not limited to blocking agents (*e.g.*, BSA or salmon or herring sperm carrier DNA), detergents (*e.g.*, SDS), chelating agents (*e.g.*, EDTA), Ficoll, PVP and the like. When using nylon membranes, in particular, an additional preferred, non-limiting example of stringent hybridization conditions is hybridization in 0.25-0.5M NaH_2PO_4, 7% SDS at about 65°C, followed by one or more washes at 0.02M NaH_2PO_4, 1% SDS at 65°C, see *e.g.*, Church and Gilbert (1984) *Proc. Natl. Acad. Sci. USA* 81:1991-1995, (or alternatively 0.2X SSC, 1% SDS).

Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1 or 3 corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein).

In addition to naturally-occurring allelic variants of the C/SKARP-1 sequences that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into the nucleotide sequences of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, thereby leading to changes in the amino acid sequence of the encoded C/SKARP-1 proteins, without altering the functional ability of the C/SKARP-1 proteins. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in the sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of C/SKARP-1 (*e.g.*, the sequence of SEQ ID NO:2) without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are conserved among the C/SKARP-1 proteins of the present invention, *e.g.*, those present in the ankyrin repeat domain, are predicted to be particularly unamenable to alteration. Furthermore, additional amino acid residues that are conserved

between the C/SKARP-1 proteins of the present invention and other ankyrin repeat containing kinases are not likely to be amenable to alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding C/SKARP-1 proteins that contain changes in amino acid residues that are not essential for activity. Such C/SKARP-1 proteins differ in amino acid sequence from SEQ ID NO:2, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to SEQ ID NO:2 (*e.g.*, to the entire length of SEQ ID NO:2).

An isolated nucleic acid molecule encoding a C/SKARP-1 protein homologous to the protein of SEQ ID NO:2 can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____ by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a C/SKARP-1 protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a C/SKARP-1 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for C/SKARP-1 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In a preferred embodiment, a mutant C/SKARP-1 protein can be assayed for the ability to 1) regulate transmission of signals from cellular receptors, *e.g.*, cardiac cell growth factor receptors; 2) modulate the entry of cells, *e.g.*, cardiac precursor cells, into mitosis; 3)

modulate cellular differentiation; 4) modulate cell death; and 5) regulate cytoskeleton function, *e.g.*, actin bundling.

In addition to the nucleic acid molecules encoding C/SKARP-1 proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. In an exemplary embodiment, the invention provides an isolated nucleic acid molecule which is antisense to a C/SKARP-1 nucleic acid molecule (*e.g.*, is antisense to the coding strand of a C/SKARP-1 nucleic acid molecule). An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire C/SKARP-1 coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding C/SKARP-1. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the coding region of human C/SKARP-1 corresponds to SEQ ID NO:3). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding C/SKARP-1. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding C/SKARP-1 disclosed herein (*e.g.*, SEQ ID NO:3), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of C/SKARP-1 mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of C/SKARP-1 mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of C/SKARP-1 mRNA (*e.g.*, between the –10 and +10 regions of the start site of a gene nucleotide sequence). An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-

carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5- oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a C/SKARP-1 protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention include direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids. Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave C/SKARP-1 mRNA transcripts to thereby inhibit translation of C/SKARP-1 mRNA. A ribozyme having specificity for a C/SKARP-1-encoding nucleic acid can be designed based upon the nucleotide sequence of a C/SKARP-1 cDNA disclosed herein (*i.e.*, SEQ ID NO:1 or 3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number _____). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a C/SKARP-1-encoding mRNA. See, *e.g.*, Cech *et al.* U.S. Patent No. 4,987,071; and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, C/SKARP-1 mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

Alternatively, C/SKARP-1 gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the C/SKARP-1 (*e.g.*, the C/SKARP-1 promoter and/or enhancers) to form triple helical structures that prevent transcription of the C/SKARP-1 gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

In yet another embodiment, the C/SKARP-1 nucleic acid molecules of the present invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, *e.g.*, the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acid molecules can be modified to generate peptide nucleic acids (see Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, *e.g.*, DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al.* Proc. Natl. Acad. Sci. 93: 14670-675.

PNAs of C/SKARP-1 nucleic acid molecules can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, for example, inducing transcription or translation arrest or inhibiting replication. PNAs of C/SKARP-1 nucleic acid molecules can

also be used in the analysis of single base pair mutations in a gene, (*e.g.*, by PNA-directed PCR clamping); as 'artificial restriction enzymes' when used in combination with other enzymes, (*e.g.*, S1 nucleases (Hyrup B. (1996) *supra*)); or as probes or primers for DNA sequencing or hybridization (Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

In another embodiment, PNAs of C/SKARP-1 can be modified, (*e.g.*, to enhance their stability or cellular uptake), by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras of C/SKARP-1 nucleic acid molecules can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, (*e.g.*, RNase H and DNA polymerases), to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup B. (1996) *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup B. (1996) *supra* and Finn P.J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs, *e.g.*, 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used as a between the PNA and the 5' end of DNA (Mag, M. *et al.* (1989) *Nucleic Acid Res.* 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn P.J. *et al.* (1996) *supra*). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (*e.g.*, for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, *e.g.*, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, *e.g.*, PCT Publication No. W089/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (See, *e.g.*, Krol *et al.* (1988) *Bio-Techniques* 6:958-976) or intercalating agents. (See, *e.g.*, Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, (*e.g.*, a peptide, hybridization triggered cross-linking agent, transport agent, or hybridization-triggered cleavage agent).

II. Isolated C/SKARP-1 Proteins and Anti-C/SKARP-1 Antibodies

One aspect of the invention pertains to isolated or recombinant C/SKARP-1 proteins and polypeptides, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise anti-C/SKARP-1 antibodies. In one embodiment,

native C/SKARP-1 proteins can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, C/SKARP-1 proteins are produced by recombinant DNA techniques. Alternative to recombinant expression, a C/SKARP-1 protein or polypeptide can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the C/SKARP-1 protein is derived, or substantially free from chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of C/SKARP-1 protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of C/SKARP-1 protein having less than about 30% (by dry weight) of non-C/SKARP-1 protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-C/SKARP-1 protein, still more preferably less than about 10% of non-C/SKARP-1 protein, and most preferably less than about 5% non-C/SKARP-1 protein. When the C/SKARP-1 protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation.

The language "substantially free of chemical precursors or other chemicals" includes preparations of C/SKARP-1 protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of C/SKARP-1 protein having less than about 30% (by dry weight) of chemical precursors or non-C/SKARP-1 chemicals, more preferably less than about 20% chemical precursors or non-C/SKARP-1 chemicals, still more preferably less than about 10% chemical precursors or non-C/SKARP-1 chemicals, and most preferably less than about 5% chemical precursors or non-C/SKARP-1 chemicals.

As used herein, a "biologically active portion" of a C/SKARP-1 protein includes a fragment of a C/SKARP-1 protein which participates in an interaction between a C/SKARP-1 molecule and a non-C/SKARP-1 molecule. Biologically active portions of a C/SKARP-1 protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the C/SKARP-1 protein, *e.g.*, the amino acid sequence shown in SEQ ID NO:2, which include less amino acids than the full length C/SKARP-1 proteins, and exhibit at least one activity of a C/SKARP-1 protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the C/SKARP-1 protein, *e.g.*, modulating signaling pathways associated with cellular growth

and differentiation. A biologically active portion of a C/SKARP-1 protein can be a polypeptide which is, for example, 10, 25, 50, 100, 200 or more amino acids in length. Biologically active portions of a C/SKARP-1 protein can be used as targets for developing agents which modulate a C/SKARP-1 mediated activity, *e.g.*, the modulation of signaling pathways associated with cellular growth and differentiation.

In one embodiment, a biologically active portion of a C/SKARP-1 protein comprises at least one ankyrin repeat domain. It is to be understood that a preferred biologically active portion of a C/SKARP-1 protein of the present invention may contain at least one ankyrin repeat domain. Another preferred biologically active portion of a C/SKARP-1 protein may contain at least one, two, three, four, five or six ankyrin repeats. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of a native C/SKARP-1 protein.

Another aspect of the invention features fragments of the protein having the amino acid sequence of SEQ ID NO:2, for example, for use as immunogens. In one embodiment, a fragment comprises at least 5 amino acids (*e.g.*, contiguous or consecutive amino acids) of the amino acid sequence of SEQ ID NO:2, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number _____. In another embodiment, a fragment comprises at least 10, 15, 20, 25, 30, 35, 40, 45, 50 or more amino acids (*e.g.*, contiguous or consecutive amino acids) of the amino acid sequence of SEQ ID NO:2, or an amino acid sequence encoded by the DNA insert of the plasmid deposited with the ATCC as Accession Number _____.

In a preferred embodiment, a C/SKARP-1 protein has an amino acid sequence shown in SEQ ID NO:2. In other embodiments, the C/SKARP-1 protein is substantially homologous to SEQ ID NO:2, and retains the functional activity of the protein of SEQ ID NO:2, yet differs in amino acid sequence due to natural allelic variation or mutagenesis, as described in detail in subsection I above. In another embodiment, the C/SKARP-1 protein is a protein which comprises an amino acid sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or more homologous to SEQ ID NO:2.

In another embodiment, the invention features a C/SKARP-1 protein which is encoded by a nucleic acid molecule consisting of a nucleotide sequence at least about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or more identical to a nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, or a complement thereof. This invention further features a C/SKARP-1 protein which is encoded by a nucleic acid molecule consisting of a nucleotide sequence which hybridizes under stringent hybridization conditions to a complement of a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1 or SEQ ID NO:3, or a complement thereof.

To determine the percent identity of two amino acid sequences or of two nucleic acid sequences, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in one or both of a first and a second amino acid or nucleic acid sequence for optimal alignment and non-homologous sequences can be disregarded for comparison purposes). In a preferred embodiment, the length of a reference sequence aligned for comparison purposes is at least 30%, preferably at least 40%, more preferably at least 50%, even more preferably at least 60%, and even more preferably at least 70%, 80%, or 90% of the length of the reference sequence (*e.g.*, when aligning a second sequence to the C/SKARP-1 amino acid sequence of SEQ ID NO:2 having 323 amino acid residues, at least 97, preferably at least 129, more preferably at least 161, even more preferably at least 194, and even more preferably at least 226, 258 or 291 amino acid residues are aligned). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position (as used herein amino acid or nucleic acid "identity" is equivalent to amino acid or nucleic acid "homology"). The percent identity between the two sequences is a function of the number of identical positions shared by the sequences, taking into account the number of gaps, and the length of each gap, which need to be introduced for optimal alignment of the two sequences.

The comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. In a preferred embodiment, the percent identity between two amino acid sequences is determined using the Needleman and Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) algorithm which has been incorporated into the GAP program in the GCG software package (available at http://www.gcg.com), using either a Blossom 62 matrix or a PAM250 matrix, and a gap weight of 16, 14, 12, 10, 8, 6, or 4 and a length weight of 1, 2, 3, 4, 5, or 6. In yet another preferred embodiment, the percent identity between two nucleotide sequences is determined using the GAP program in the GCG software package (available at http://www.gcg.com), using a NWSgapdna.CMP matrix and a gap weight of 40, 50, 60, 70, or 80 and a length weight of 1, 2, 3, 4, 5, or 6. A preferred, non-limiting example of parameters to be used in conjunction with the GAP program include a Blosum 62 scoring matrix with a gap penalty of 12, a gap extend penalty of 4, and a frameshift gap penalty of 5.

In another embodiment, the percent identity between two amino acid or nucleotide sequences is determined using the algorithm of E. Meyers and W. Miller (CABIOS, 4:11-17 (1989)) which has been incorporated into the ALIGN program (version 2.0 or version 2.U), using a PAM120 weight residue table, a gap length penalty of 12 and a gap penalty of 4.

The nucleic acid and protein sequences of the present invention can further be used as a "query sequence" to perform a search against public databases to, for example, identify

other family members or related sequences. Such searches can be performed using the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to C/SKARP-1 nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to C/SKARP-1 protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (*e.g.*, XBLAST and NBLAST) can be used. See http://www.ncbi.nlm.nih.gov.

The invention also provides C/SKARP-1 chimeric or fusion proteins. As used herein, a C/SKARP-1 "chimeric protein" or "fusion protein" comprises a C/SKARP-1 polypeptide operatively linked to a non-C/SKARP-1 polypeptide. An "C/SKARP-1 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to C/SKARP-1, whereas a "non-C/SKARP-1 polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the C/SKARP-1 protein, *e.g.*, a protein which is different from the C/SKARP-1 protein and which is derived from the same or a different organism. Within a C/SKARP-1 fusion protein the C/SKARP-1 polypeptide can correspond to all or a portion of a C/SKARP-1 protein. In a preferred embodiment, a C/SKARP-1 fusion protein comprises at least one biologically active portion of a C/SKARP-1 protein. In another preferred embodiment, a C/SKARP-1 fusion protein comprises at least two biologically active portions of a C/SKARP-1 protein. Within the fusion protein, the term "operatively linked" is intended to indicate that the C/SKARP-1 polypeptide and the non-C/SKARP-1 polypeptide are fused in-frame to each other. The non-C/SKARP-1 polypeptide can be fused to the N-terminus or C-terminus of the C/SKARP-1 polypeptide.

For example, in one embodiment, the fusion protein is a GST-C/SKARP-1 fusion protein in which the C/SKARP-1 sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant C/SKARP-1.

In another embodiment, the fusion protein is a C/SKARP-1 protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of C/SKARP-1 can be increased through use of a heterologous signal sequence.

The C/SKARP-1 fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject *in vivo*. The C/SKARP-1 fusion proteins can be used to affect the bioavailability of a C/SKARP-1 substrate. Use of C/SKARP-1 fusion proteins may be useful therapeutically for the treatment of disorders

caused by, for example, (i) aberrant modification or mutation of a gene encoding a C/SKARP-1 protein; (ii) mis-regulation of the C/SKARP-1 gene; and (iii) aberrant post-translational modification of a C/SKARP-1 protein.

Moreover, the C/SKARP-1-fusion proteins of the invention can be used as immunogens to produce anti-C/SKARP-1 antibodies in a subject, to purify C/SKARP-1 ligands and in screening assays to identify molecules which inhibit the interaction of C/SKARP-1 with a C/SKARP-1 substrate.

Preferably, a C/SKARP-1 chimeric or fusion protein of the invention is produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). A C/SKARP-1-encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the C/SKARP-1 protein.

The present invention also pertains to variants of the C/SKARP-1 proteins which function as either C/SKARP-1 agonists (mimetics) or as C/SKARP-1 antagonists. Variants of the C/SKARP-1 proteins can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of a C/SKARP-1 protein. An agonist of the C/SKARP-1 proteins can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of a C/SKARP-1 protein. An antagonist of a C/SKARP-1 protein can inhibit one or more of the activities of the naturally occurring form of the C/SKARP-1 protein by, for example, competitively modulating a C/SKARP-1-mediated activity of a C/SKARP-1 protein. Thus, specific biological effects can be elicited by treatment with a variant of limited function. In one embodiment, treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein has fewer side effects in a subject relative to treatment with the naturally occurring form of the C/SKARP-1 protein.

In one embodiment, variants of a C/SKARP-1 protein which function as either C/SKARP-1 agonists (mimetics) or as C/SKARP-1 antagonists can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of a C/SKARP-1

protein for C/SKARP-1 protein agonist or antagonist activity. In one embodiment, a variegated library of C/SKARP-1 variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of C/SKARP-1 variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential C/SKARP-1 sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of C/SKARP-1 sequences therein. There are a variety of methods which can be used to produce libraries of potential C/SKARP-1 variants from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential C/SKARP-1 sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477.

In addition, libraries of fragments of a C/SKARP-1 protein coding sequence can be used to generate a variegated population of C/SKARP-1 fragments for screening and subsequent selection of variants of a C/SKARP-1 protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of a C/SKARP-1 coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the C/SKARP-1 protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of C/SKARP-1 proteins. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recrusive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify

C/SKARP-1 variants (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

In one embodiment, cell based assays can be exploited to analyze a variegated C/SKARP-1 library. For example, a library of expression vectors can be transfected into a cell line, *e.g.*, a cardiac cell line, which ordinarily responds to a particular ligand in a C/SKARP-1-dependent manner. The transfected cells are then contacted with the ligand and the effect of expression of the mutant on signaling by the ligand can be detected, *e.g.*, by monitoring intracellular calcium, IP3, or diacylglycerol concentration, phosphorylation profile of intracellular proteins, cell proliferation and/or migration, or the activity of a C/SKARP-1-regulated transcription factor. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of signaling by the ligand, and the individual clones further characterized.

An isolated C/SKARP-1 protein, or a portion or fragment thereof, can be used as an immunogen to generate antibodies that bind C/SKARP-1 using standard techniques for polyclonal and monoclonal antibody preparation. A full-length C/SKARP-1 protein can be used or, alternatively, the invention provides antigenic peptide fragments of C/SKARP-1 for use as immunogens. The antigenic peptide of C/SKARP-1 comprises at least 8 amino acid residues of the amino acid sequence shown in SEQ ID NO:2 and encompasses an epitope of C/SKARP-1 such that an antibody raised against the peptide forms a specific immune complex with C/SKARP-1. Preferably, the antigenic peptide comprises at least 10 amino acid residues, more preferably at least 15 amino acid residues, even more preferably at least 20 amino acid residues, and most preferably at least 30 amino acid residues.

Preferred epitopes encompassed by the antigenic peptide are regions of C/SKARP-1 that are located on the surface of the protein, *e.g.*, hydrophilic regions, as well as regions with high antigenicity (see, for example, Figure 3).

A C/SKARP-1 immunogen typically is used to prepare antibodies by immunizing a suitable subject, (*e.g.*, rabbit, goat, mouse or other mammal) with the immunogen. An appropriate immunogenic preparation can contain, for example, recombinantly expressed C/SKARP-1 protein or a chemically synthesized C/SKARP-1 polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent. Immunization of a suitable subject with an immunogenic C/SKARP-1 preparation induces a polyclonal anti-C/SKARP-1 antibody response.

Accordingly, another aspect of the invention pertains to anti-C/SKARP-1 antibodies. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, *i.e.*, molecules that contain an antigen binding site which specifically binds (immunoreacts with) an antigen, such as C/SKARP-1. Examples of immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an

enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies that bind C/SKARP-1. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope of C/SKARP-1. A monoclonal antibody composition thus typically displays a single binding affinity for a particular C/SKARP-1 protein with which it immunoreacts.

Polyclonal anti-C/SKARP-1 antibodies can be prepared as described above by immunizing a suitable subject with a C/SKARP-1 immunogen. The anti-C/SKARP-1 antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized C/SKARP-1. If desired, the antibody molecules directed against C/SKARP-1 can be isolated from the mammal (*e.g.*, from the blood) and further purified by well known techniques, such as protein A chromatography to obtain the IgG fraction. At an appropriate time after immunization, *e.g.*, when the anti-C/SKARP-1 antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497) (see also, Brown *et al.* (1981) *J. Immunol.* 127:539-46; Brown *et al.* (1980) *J. Biol. Chem* .255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; and Yeh *et al.* (1982) *Int. J. Cancer* 29:269-75), the more recent human B cell hybridoma technique (Kozbor *et al.* (1983) *Immunol Today* 4:72), the EBV-hybridoma technique (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing monoclonal antibody hybridomas is well known (see generally R. H. Kenneth, in *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, New York (1980); E. A. Lerner (1981) *Yale J. Biol. Med.*, 54:387-402; M. L. Gefter *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Briefly, an immortal cell line (typically a myeloma) is fused to lymphocytes (typically splenocytes) from a mammal immunized with a C/SKARP-1 immunogen as described above, and the culture supernatants of the resulting hybridoma cells are screened to identify a hybridoma producing a monoclonal antibody that binds C/SKARP-1.

Any of the many well known protocols used for fusing lymphocytes and immortalized cell lines can be applied for the purpose of generating an anti-C/SKARP-1 monoclonal antibody (see, *e.g.*, G. Galfre *et al.* (1977) *Nature* 266:55052; Gefter *et al.* *Somatic Cell Genet.*, cited *supra*; Lerner, *Yale J. Biol. Med.*, cited *supra*; Kenneth, *Monoclonal Antibodies*, cited *supra*). Moreover, the ordinarily skilled worker will appreciate that there are many variations of such methods which also would be useful. Typically, the immortal cell line (*e.g.*, a myeloma cell line) is derived from the same mammalian species as the lymphocytes. For example, murine hybridomas can be made by

fusing lymphocytes from a mouse immunized with an immunogenic preparation of the present invention with an immortalized mouse cell line. Preferred immortal cell lines are mouse myeloma cell lines that are sensitive to culture medium containing hypoxanthine, aminopterin and thymidine ("HAT medium"). Any of a number of myeloma cell lines can be used as a fusion partner according to standard techniques, *e.g.*, the P3-NS1/1-Ag4-1, P3-x63-Ag8.653 or Sp2/O-Ag14 myeloma lines. These myeloma lines are available from ATCC. Typically, HAT-sensitive mouse myeloma cells are fused to mouse splenocytes using polyethylene glycol ("PEG"). Hybridoma cells resulting from the fusion are then selected using HAT medium, which kills unfused and unproductively fused myeloma cells (unfused splenocytes die after several days because they are not transformed). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind C/SKARP-1, *e.g.*, using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal anti-C/SKARP-1 antibody can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (*e.g.*, an antibody phage display library) with C/SKARP-1 to thereby isolate immunoglobulin library members that bind C/SKARP-1. Kits for generating and screening phage display libraries are commercially available (*e.g.*, the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, Ladner *et al.* U.S. Patent No. 5,223,409; Kang *et al.* PCT International Publication No. WO 92/18619; Dower *et al.* PCT International Publication No. WO 91/17271; Winter *et al.* PCT International Publication WO 92/20791; Markland *et al.* PCT International Publication No. WO 92/15679; Breitling *et al.* PCT International Publication WO 93/01288; McCafferty *et al.* PCT International Publication No. WO 92/01047; Garrard *et al.* PCT International Publication No. WO 92/09690; Ladner *et al.* PCT International Publication No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol. Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrad *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acid Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; and McCafferty *et al. Nature* (1990) 348:552-554.

Additionally, recombinant anti-C/SKARP-1 antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. Such chimeric and humanized monoclonal antibodies can be produced by

recombinant DNA techniques known in the art, for example using methods described in Robinson *et al.* International Application No. PCT/US86/02269; Akira, *et al.* European Patent Application 184,187; Taniguchi, M., European Patent Application 171,496; Morrison *et al.* European Patent Application 173,494; Neuberger *et al.* PCT International Publication No. WO 86/01533; Cabilly *et al.* U.S. Patent No. 4,816,567; Cabilly *et al.* European Patent Application 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Canc. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; and Shaw *et al.* (1988) *J. Natl. Cancer Inst.* 80:1553-1559); Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *BioTechniques* 4:214; Winter U.S. Patent 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeyan *et al.* (1988) *Science* 239:1534; and Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

An anti-C/SKARP-1 antibody (*e.g.*, monoclonal antibody) can be used to isolate C/SKARP-1 by standard techniques, such as affinity chromatography or immunoprecipitation. An anti-C/SKARP-1 antibody can facilitate the purification of natural C/SKARP-1 from cells and of recombinantly produced C/SKARP-1 expressed in host cells. Moreover, an anti-C/SKARP-1 antibody can be used to detect C/SKARP-1 protein (*e.g.*, in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the C/SKARP-1 protein. Anti-C/SKARP-1 antibodies can be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, *e.g.*, to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling (*i.e.*, physically linking) the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, -galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, for example recombinant expression vectors, containing a C/SKARP-1 nucleic acid molecule or vectors containing a nucleic acid molecule which encodes a C/SKARP-1 protein (or a portion thereof). As used

herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cells and those which direct expression of the nucleotide sequence only in certain host cells (*e.g.*, tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, and the like. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, C/SKARP-1 proteins, mutant forms of C/SKARP-1 proteins, fusion proteins, and the like).

Accordingly, an exemplary embodiment provides a method for producing a protein, preferably a C/SKARP-1 protein, by culturing in a suitable medium a host cell of the invention (*e.g.*, a mammalian host cell such as a non-human mammalian cell) containing a recombinant expression vector, such that the protein is produced.

The recombinant expression vectors of the invention can be designed for expression of C/SKARP-1 proteins in prokaryotic or eukaryotic cells. For example, C/SKARP-1 proteins can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Purified fusion proteins can be utilized in C/SKARP-1 activity assays, (*e.g.*, direct assays or competitive assays described in detail below), or to generate antibodies specific for C/SKARP-1 proteins, for example. In a preferred embodiment, a C/SKARP-1 fusion protein expressed in a retroviral expression vector of the present invention can be utilized to infect bone marrow cells which are subsequently transplanted into irradiated recipients. The pathology of the subject recipient is then examined after sufficient time has passed (*e.g.*, six (6) weeks).

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase

transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the C/SKARP-1 expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerivisae* include pYepSec1 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and picZ (InVitrogen Corp, San Diego, CA).

Alternatively, C/SKARP-1 proteins can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (*e.g.*, tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in

particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (*e.g.*, the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (*e.g.*, milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to C/SKARP-1 mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) 1986.

Another aspect of the invention pertains to host cells into which a C/SKARP-1 nucleic acid molecule of the invention is introduced, *e.g.*, a C/SKARP-1 nucleic acid molecule within a vector (*e.g.*, a recombinant expression vector) or a C/SKARP-1 nucleic acid molecule containing sequences which allow it to homologously recombine into a specific site of the host cell's genome. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, a C/SKARP-1 protein can be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those skilled in the art.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding a C/SKARP-1 protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) a C/SKARP-1 protein. Accordingly, the invention further provides methods for producing a C/SKARP-1 protein using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of the invention (into which a recombinant expression vector encoding a C/SKARP-1 protein has been introduced) in a suitable medium such that a C/SKARP-1 protein is produced. In another embodiment, the method further comprises isolating a C/SKARP-1 protein from the medium or the host cell.

The host cells of the invention can also be used to produce non-human transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which C/SKARP-1-coding sequences have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous C/SKARP-1 sequences have been introduced into their genome or homologous recombinant animals in which endogenous C/SKARP-1 sequences have been altered. Such animals are useful for studying the function and/or activity of a C/SKARP-1 and for identifying and/or evaluating modulators of C/SKARP-1 activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene.

Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, and the like. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, a "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous C/SKARP-1 gene has been altered by homologous recombination between the endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, *e.g.*, an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing a C/SKARP-1-encoding nucleic acid into the male pronuclei of a fertilized oocyte, *e.g.*, by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. The C/SKARP-1 cDNA sequence of SEQ ID NO:1 can be introduced as a transgene into the genome of a non-human animal. Alternatively, a nonhuman homologue of a human C/SKARP-1 gene, such as a mouse or rat C/SKARP-1 gene, can be used as a transgene. Alternatively, a C/SKARP-1 gene homologue, such as another C/SKARP-1 family member, can be isolated based on hybridization to the C/SKARP-1 cDNA sequences of SEQ ID NO:1 or 3, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____ (described further in subsection I above) and used as a transgene. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to a C/SKARP-1 transgene to direct expression of a C/SKARP-1 protein to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009, both by Leder *et al.*, U.S. Patent No. 4,873,191 by Wagner *et al.* and in Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of a C/SKARP-1 transgene in its genome and/or expression of C/SKARP-1 mRNA in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying a transgene encoding a C/SKARP-1 protein can further be bred to other transgenic animals carrying other transgenes.

To create a homologous recombinant animal, a vector is prepared which contains at least a portion of a C/SKARP-1 gene into which a deletion, addition or substitution has been introduced to thereby alter, *e.g.*, functionally disrupt, the C/SKARP-1 gene. The C/SKARP-1 gene can be a human gene (*e.g.*, the cDNA of SEQ ID NO:3), but more preferably, is a

non-human homologue of a human C/SKARP-1 gene (*e.g.*, a cDNA isolated by stringent hybridization with the nucleotide sequence of SEQ ID NO:1). For example, a mouse C/SKARP-1 gene can be used to construct a homologous recombination nucleic acid molecule, *e.g.*, a vector, suitable for altering an endogenous C/SKARP-1 gene in the mouse genome. In a preferred embodiment, the homologous recombination nucleic acid molecule is designed such that, upon homologous recombination, the endogenous C/SKARP-1 gene is functionally disrupted (*i.e.*, no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the homologous recombination nucleic acid molecule can be designed such that, upon homologous recombination, the endogenous C/SKARP-1 gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous C/SKARP-1 protein). In the homologous recombination nucleic acid molecule, the altered portion of the C/SKARP-1 gene is flanked at its 5' and 3' ends by additional nucleic acid sequence of the C/SKARP-1 gene to allow for homologous recombination to occur between the exogenous C/SKARP-1 gene carried by the homologous recombination nucleic acid molecule and an endogenous C/SKARP-1 gene in a cell, *e.g.*, an embryonic stem cell. The additional flanking C/SKARP-1 nucleic acid sequence is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the homologous recombination nucleic acid molecule (see, *e.g.*, Thomas, K.R. and Capecchi, M. R. (1987) *Cell* 51:503 for a description of homologous recombination vectors). The homologous recombination nucleic acid molecule is introduced into a cell, *e.g.*, an embryonic stem cell line (*e.g.*, by electroporation) and cells in which the introduced C/SKARP-1 gene has homologously recombined with the endogenous C/SKARP-1 gene are selected (see *e.g.*, Li, E. *et al.* (1992) *Cell* 69:915). The selected cells can then injected into a blastocyst of an animal (*e.g.*, a mouse) to form aggregation chimeras (see *e.g.*, Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E.J. Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination nucleic acid molecules, *e.g.*, vectors, or homologous recombinant animals are described further in Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 and in PCT International Publication Nos.: WO 90/11354 by Le Mouellec *et al.*; WO 91/01140 by Smithies *et al.*; WO 92/0968 by Zijlstra *et al.*; and WO 93/04169 by Berns *et al.*

In another embodiment, transgenic non-humans animals can be produced which contain selected systems which allow for regulated expression of the transgene. One

example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, see, *e.g.*, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, *e.g.*, by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut, I. *et al.* (1997) *Nature* 385:810-813 and PCT International Publication Nos. WO 97/07668 and WO 97/07669. In brief, a cell, *e.g.*, a somatic cell, from the transgenic animal can be isolated and induced to exit the growth cycle and enter G_O phase. The quiescent cell can then be fused, *e.g.*, through the use of electrical pulses, to an enucleated oocyte from an animal of the same species from which the quiescent cell is isolated. The reconstructed oocyte is then cultured such that it develops to morula or blastocyte and then transferred to pseudopregnant female foster animal. The offspring borne of this female foster animal will be a clone of the animal from which the cell, *e.g.*, the somatic cell, is isolated.

IV. Pharmaceutical Compositions

The C/SKARP-1 nucleic acid molecules, C/SKARP-1 proteins, fragments thereof, anti-C/SKARP-1 antibodies, and C/SKARP-1 modulators (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components:

a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL™ (BASF, Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as manitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (*e.g.*, a C/SKARP-1 protein, nucleic acid molecule, anti-C/SKARP-1 antibody, or C/SKARP-1 modulators) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed. Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from pressured container or dispenser which contains a suitable propellant, *e.g.*, a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (*e.g.*, with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds which exhibit large therapeutic indices are preferred. While compounds that exhibit toxic side effects may be used, care should be taken to design a delivery system that targets such compounds to the site of affected tissue in order to minimize potential damage to uninfected cells and, thereby, reduce side effects.

The data obtained from the cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose may be formulated in animal models to achieve a circulating plasma concentration range that includes the IC50 (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans. Levels in plasma may be measured, for example, by high performance liquid chromatography.

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (see U.S. Patent 5,328,470) or by stereotactic injection (see *e.g.*, Chen *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, *e.g.*, retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, antibodies and modulators described herein can be used in one or more of the following methods: a) screening assays; b) predictive medicine (*e.g.*, diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenetics); and c) methods of treatment (*e.g.*, therapeutic and prophylactic). As described herein, a C/SKARP-1 protein of the invention has one or more of the following activities: (i) mediation of specific macromoleculer interactions; (ii) mediation of interactions between proteins and/or between regions of a single protein; (iii) formation of binding sites for distinct proteins (*e.g.*, non-C/SKARP proteins); (iv) bridging of cellular components; (v) regulation of gene expression (*e.g.*, cardiac gene expression) and, thus, can be used to, for example, (1) modulate cellular localization (*e.g.*, anchoring C/SKARP binding proteins in a specific cellular localization); (2) modulate development and/or differentiation (*e.g.*, myogenic development and/or differentiation, heart development and/or differentiation); (3) modulate cardiac maturation and/or morphogenesis; (4) as a marker (*e.g.*, an early marker) of cardiac and/or myogenic cell lineage; and (5) modulate and/or treat C/SKARP-1-associated or related disorders.

As used herein, a "C/SKARP-1-associated or related disorder" includes a disorder, disease, or condition which is caused or characterized by a misregulation (*e.g.*, downregulation or upregulation) of C/SKARP-1 activity. The C/SKARP-1 molecules of the present invention may also act as novel diagnostic targets and therapeutic agents for cardiovascular diseases or disorders. Exemplary C/SKARP-related disorders include, but are not limited to, cardiac hypertrophy, cardiac disorders and/or cardiovascular disease (*e.g.*, congestive heart failure, cardiomyopathy and the like. Additional exemplary C/SKARP-1-associated disorders include, but are not limited to disorders such as arteriosclerosis, ischemia reperfusion injury, restenosis, arterial inflammation, vascular wall remodeling, ventricular remodeling, rapid ventricular pacing, coronary microembolism, tachycardia, bradycardia, pressure overload, aortic bending, coronary artery ligation, vascular heart disease, atrial fibrillation, long-QT syndrome, congestive heart failure, sinus node dysfunction, angina, heart failure, hypertension, atrial fibrillation, atrial flutter, dilated cardiomyopathy, idiopathic cardiomyopathy, myocardial infarction, coronary artery disease, coronary artery spasm, ischemic disease, arrhythmia, and cardiovascular developmental disorders (*e.g.*, arteriovenous malformations, arteriovenous fistulae, raynaud's syndrome, neurogenic thoracic outlet syndrome, causalgia/reflex sympathetic dystrophy, hemangioma, aneurysm, cavernous angioma, aortic valve stenosis, atrial septal defects, atrioventricular canal, coarctation of the aorta, ebsteins anomaly, hypoplastic left heart syndrome,

interruption of the aortic arch, mitral valve prolapse, ductus arteriosus, patent foramen ovale, partial anomalous pulmonary venous return, pulmonary atresia with ventricular septal defect, pulmonary atresia without ventricular septal defect, persistance of the fetal circulation, pulmonary valve stenosis, single ventricle, total anomalous pulmonary venous return, transposition of the great vessels, tricuspid atresia, truncus arteriosus, ventricular septal defects). A cardiovasular disease or disorder also includes an endothelial cell disorder. As used herein, an "endothelial cell disorder" includes a disorder characterized by aberrant, unregulated, or unwanted endothelial cell activity, *e.g.*, proliferation, migration, angiogenesis, or vascularization; or aberrant expression of cell surface adhesion molecules or genes associated with angiogenesis, *e.g.*, TIE-2, FLT and FLK. Endothelial cell disorders include tumorigenesis, tumor metastasis, psoriasis, diabetic retinopathy, endometriosis, Grave's disease, ischemic disease (*e.g.*, atherosclerosis), and chronic inflammatory diseases (*e.g.*, rheumatoid arthritis).

The isolated nucleic acid molecules of the invention can be used, for example, to express C/SKARP-1 protein (*e.g.*, *via* a recombinant expression vector in a host cell in gene therapy applications), to detect C/SKARP-1 mRNA (*e.g.*, in a biological sample) or a genetic alteration in a C/SKARP-1 gene, and to modulate C/SKARP-1 activity, as described further below. The C/SKARP-1 proteins can be used to treat disorders characterized by insufficient or excessive production of a C/SKARP-1 substrate or production of C/SKARP-1 inhibitors. In addition, the C/SKARP-1 proteins can be used to screen for naturally occurring C/SKARP-1 substrates, to screen for drugs or compounds which modulate C/SKARP-1 activity, as well as to treat disorders characterized by insufficient or excessive production of C/SKARP-1 protein or production of C/SKARP-1 protein forms which have decreased, aberrant or unwanted activity compared to C/SKARP-1 wild type protein (*e.g.*, C/SKARP-1-associated disorders). Moreover, the anti-C/SKARP-1 antibodies of the invention can be used to detect and isolate C/SKARP-1 proteins, regulate the bioavailability of C/SKARP-1 proteins, and modulate C/SKARP-1 activity.

A. Screening Assays:

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, *i.e.*, candidate or test compounds or agents (*e.g.*, peptides, peptidomimetics, small molecules or other drugs) which bind to C/SKARP-1 proteins, have a stimulatory or inhibitory effect on, for example, C/SKARP-1 expression or C/SKARP-1 activity, or have a stimulatory or inhibitory effect on, for example, the expression or activity of C/SKARP-1 substrate.

In one embodiment, the invention provides assays for screening candidate or test compounds which are substrates of a C/SKARP-1 protein or polypeptide or biologically active portion thereof. In another embodiment, the invention provides assays for screening

candidate or test compounds which bind to or modulate the activity of a C/SKARP-1 protein or polypeptide or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the 'one-bead one-compound' library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam, K.S. (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt *et al.* (1993) *Proc. Natl. Acad. Sci. U.S.A.* 90:6909; Erb *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann *et al.* (1994). *J. Med. Chem.* 37:2678; Cho *et al.* (1993) *Science* 261:1303; Carrell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2059; Carell *et al.* (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and in Gallop *et al.* (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (*e.g.*, Houghten (1992) *Biotechniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (Ladner USP 5,223,409), spores (Ladner USP '409), plasmids (Cull *et al.* (1992) *Proc Natl Acad Sci USA* 89:1865-1869) or on phage (Scott and Smith (1990) *Science* 249:386-390); (Devlin (1990) *Science* 249:404-406); (Cwirla *et al.* (1990) *Proc. Natl. Acad. Sci.* 87:6378-6382); (Felici (1991) *J. Mol. Biol.* 222:301-310); (Ladner *supra.*).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a C/SKARP-1 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate C/SKARP-1 activity is determined. Determining the ability of the test compound to modulate C/SKARP-1 activity can be accomplished by monitoring, for example, intracellular calcium, IP3, or diacylglycerol concentration, phosphorylation profile of intracellular proteins, cell proliferation and/or migration, or the activity of a C/SKARP-1-regulated transcription factor. The cell, for example, can be of mammalian origin, *e.g.*, a cardiac cell.

The ability of the test compound to modulate C/SKARP-1 binding to a substrate or to bind to C/SKARP-1 can also be determined. Determining the ability of the test compound to modulate C/SKARP-1 binding to a substrate can be accomplished, for example, by coupling the C/SKARP-1 substrate with a radioisotope or enzymatic label such that binding of the C/SKARP-1 substrate to C/SKARP-1 can be determined by detecting the labeled C/SKARP-1 substrate in a complex. Determining the ability of the test compound to bind C/SKARP-1 can be accomplished, for example, by coupling the compound with a radioisotope or enzymatic label such that binding of the compound to C/SKARP-1 can be

determined by detecting the labeled C/SKARP-1 compound in a complex. For example, compounds (*e.g.*, C/SKARP-1 substrates) can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product.

It is also within the scope of this invention to determine the ability of a compound (*e.g.*, a C/SKARP-1 substrate) to interact with C/SKARP-1 without the labeling of any of the interactants. For example, a microphysiometer can be used to detect the interaction of a compound with C/SKARP-1 without the labeling of either the compound or the C/SKARP-1. McConnell, H. M. *et al.* (1992) *Science* 257:1906-1912. As used herein, a "microphysiometer" (*e.g.*, Cytosensor) is an analytical instrument that measures the rate at which a cell acidifies its environment using a light-addressable potentiometric sensor (LAPS). Changes in this acidification rate can be used as an indicator of the interaction between a compound and C/SKARP-1.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a C/SKARP-1 target molecule (*e.g.*, a C/SKARP-1 substrate) with a test compound and determining the ability of the test compound to modulate (*e.g.* stimulate or inhibit) the activity of the C/SKARP-1 target molecule. Determining the ability of the test compound to modulate the activity of a C/SKARP-1 target molecule can be accomplished, for example, by determining the ability of the C/SKARP-1 protein to bind to or interact with the C/SKARP-1 target molecule.

Determining the ability of the C/SKARP-1 protein or a biologically active fragment thereof, to bind to or interact with a C/SKARP-1 target molecule can be accomplished by one of the methods described above for determining direct binding. In a preferred embodiment, determining the ability of the C/SKARP-1 protein to bind to or interact with a C/SKARP-1 target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (*i.e.*, intracellular Ca^{2+}, diacylglycerol, IP_3, and the like), detecting catalytic/enzymatic activity of the target an appropriate substrate, detecting the induction of a reporter gene (comprising a target-responsive regulatory element operatively linked to a nucleic acid encoding a detectable marker, *e.g.*, luciferase), or detecting a target-regulated cellular response.

In yet another embodiment, an assay of the present invention is a cell-free assay in which a C/SKARP-1 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to bind to the C/SKARP-1 protein or biologically active portion thereof is determined. Preferred biologically active portions of the C/SKARP-1 proteins to be used in assays of the present invention include fragments

which participate in interactions with non-C/SKARP-1 molecules, *e.g.*, fragments with high surface probability scores (see, for example, Figure 4). Binding of the test compound to the C/SKARP-1 protein can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the C/SKARP-1 protein or biologically active portion thereof with a known compound which binds C/SKARP-1 to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with a C/SKARP-1 protein, wherein determining the ability of the test compound to interact with a C/SKARP-1 protein comprises determining the ability of the test compound to preferentially bind to C/SKARP-1 or biologically active portion thereof as compared to the known compound.

In another embodiment, the assay is a cell-free assay in which a C/SKARP-1 protein or biologically active portion thereof is contacted with a test compound and the ability of the test compound to modulate (*e.g.*, stimulate or inhibit) the activity of the C/SKARP-1 protein or biologically active portion thereof is determined. Determining the ability of the test compound to modulate the activity of a C/SKARP-1 protein can be accomplished, for example, by determining the ability of the C/SKARP-1 protein to bind to a C/SKARP-1 target molecule by one of the methods described above for determining direct binding. Determining the ability of the C/SKARP-1 protein to bind to a C/SKARP-1 target molecule can also be accomplished using a technology such as real-time Biomolecular Interaction Analysis (BIA). Sjolander, S. and Urbaniczky, C. (1991) *Anal. Chem.* 63:2338-2345 and Szabo *et al.* (1995) *Curr. Opin. Struct. Biol.* 5:699-705. As used herein, "BIA" is a technology for studying biospecific interactions in real time, without labeling any of the interactants (*e.g.*, BIAcore). Changes in the optical phenomenon of surface plasmon resonance (SPR) can be used as an indication of real-time reactions between biological molecules.

In an alternative embodiment, determining the ability of the test compound to modulate the activity of a C/SKARP-1 protein can be accomplished by determining the ability of the C/SKARP-1 protein to further modulate the activity of a downstream effector of a C/SKARP-1 target molecule. For example, the activity of the effector molecule on an appropriate target can be determined or the binding of the effector to an appropriate target can be determined as previously described.

In yet another embodiment, the cell-free assay involves contacting a C/SKARP-1 protein or biologically active portion thereof with a known compound which binds the C/SKARP-1 protein to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the C/SKARP-1 protein, wherein determining the ability of the test compound to interact with the C/SKARP-1 protein comprises determining the ability of the C/SKARP-1 protein to preferentially bind to or modulate the activity of a C/SKARP-1 target molecule.

The cell-free assays of the present invention are amenable to use of both soluble and/or membrane-bound forms of isolated proteins (*e.g.*, C/SKARP-1 proteins or biologically active portions thereof). In the case of cell-free assays in which a membrane-bound form of an isolated protein is used it may be desirable to utilize a solubilizing agent such that the membrane-bound form of the isolated protein is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-dodecylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton® X-100, Triton® X-114, Thesit®, Isotridecypoly(ethylene glycol ether)$_n$, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either C/SKARP-1 or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the proteins, as well as to accommodate automation of the assay. Binding of a test compound to a C/SKARP-1 protein, or interaction of a C/SKARP-1 protein with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtiter plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase/ C/SKARP-1 fusion proteins or glutathione-S-transferase/target fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtiter plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or C/SKARP-1 protein, and the mixture incubated under conditions conducive to complex formation (*e.g.*, at physiological conditions for salt and pH). Following incubation, the beads or microtiter plate wells are washed to remove any unbound components, the matrix immobilized in the case of beads, complex determined either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of C/SKARP-1 binding or activity determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either a C/SKARP-1 protein or a C/SKARP-1 target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated C/SKARP-1 protein or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques known in the art (*e.g.*, biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with C/SKARP-1 protein or target molecules but which do not interfere with binding of the C/SKARP-1 protein to its

target molecule can be derivatized to the wells of the plate, and unbound target or C/SKARP-1 protein trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the C/SKARP-1 protein or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the C/SKARP-1 protein or target molecule.

In another embodiment, modulators of C/SKARP-1 expression are identified in a method wherein a cell is contacted with a candidate compound and the expression of C/SKARP-1 mRNA or protein in the cell is determined. The level of expression of C/SKARP-1 mRNA or protein in the presence of the candidate compound is compared to the level of expression of C/SKARP-1 mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of C/SKARP-1 expression based on this comparison. For example, when expression of C/SKARP-1 mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of C/SKARP-1 mRNA or protein expression. Alternatively, when expression of C/SKARP-1 mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of C/SKARP-1 mRNA or protein expression. The level of C/SKARP-1 mRNA or protein expression in the cells can be determined by methods described herein for detecting C/SKARP-1 mRNA or protein.

In yet another aspect of the invention, the C/SKARP-1 proteins can be used as "bait proteins" in a two-hybrid assay or three-hybrid assay (see, *e.g.*, U.S. Patent No. 5,283,317; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696; and Brent WO94/10300), to identify other proteins, which bind to or interact with C/SKARP-1 ("C/SKARP-1-binding proteins" or "C/SKARP-1-bp") and are involved in C/SKARP-1 activity. Such C/SKARP-1-binding proteins are also likely to be involved in the propagation of signals by the C/SKARP-1 proteins or C/SKARP-1 targets as, for example, downstream elements of a C/SKARP-1-mediated signaling pathway. Alternatively, such C/SKARP-1-binding proteins are likely to be C/SKARP-1 inhibitors.

The two-hybrid system is based on the modular nature of most transcription factors, which consist of separable DNA-binding and activation domains. Briefly, the assay utilizes two different DNA constructs. In one construct, the gene that codes for a C/SKARP-1 protein is fused to a gene encoding the DNA binding domain of a known transcription factor (*e.g.*, GAL-4). In the other construct, a DNA sequence, from a library of DNA sequences, that encodes an unidentified protein ("prey" or "sample") is fused to a gene that codes for the activation domain of the known transcription factor. If the "bait" and the "prey" proteins are

able to interact, *in vivo,* forming a C/SKARP-1-dependent complex, the DNA-binding and activation domains of the transcription factor are brought into close proximity. This proximity allows transcription of a reporter gene (*e.g.*, LacZ) which is operably linked to a transcriptional regulatory site responsive to the transcription factor. Expression of the reporter gene can be detected and cell colonies containing the functional transcription factor can be isolated and used to obtain the cloned gene which encodes the protein which interacts with the C/SKARP-1 protein.

In another aspect, the invention pertains to a combination of two or more of the assays described herein. For example, a modulating agent can be identified using a cell-based or a cell free assay, and the ability of the agent to modulate the activity of a C/SKARP-1 protein can be confirmed *in vivo*, *e.g.*, in an animal such as an animal model for a cardiovascular disorder.

This invention further pertains to novel agents identified by the above-described screening assays. Accordingly, it is within the scope of this invention to further use an agent identified as described herein in an appropriate animal model. For example, an agent identified as described herein (*e.g.*, a C/SKARP-1 modulating agent, an antisense C/SKARP-1 nucleic acid molecule, a C/SKARP-1-specific antibody, or a C/SKARP-1-binding partner) can be used in an animal model to determine the efficacy, toxicity, or side effects of treatment with such an agent. Alternatively, an agent identified as described herein can be used in an animal model to determine the mechanism of action of such an agent. Furthermore, this invention pertains to uses of novel agents identified by the above-described screening assays for treatments as described herein.

B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome; and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

1. Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. This process is called chromosome mapping. Accordingly, portions or fragments of the C/SKARP-1 nucleotide sequences, described herein, can be used to map the location of the C/SKARP-1

genes on a chromosome. The mapping of the C/SKARP-1 sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, C/SKARP-1 genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the C/SKARP-1 nucleotide sequences. Computer analysis of the C/SKARP-1 sequences can be used to predict primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the C/SKARP-1 sequences will yield an amplified fragment.

Somatic cell hybrids are prepared by fusing somatic cells from different mammals (*e.g.*, human and mouse cells). As hybrids of human and mouse cells grow and divide, they gradually lose human chromosomes in random order, but retain the mouse chromosomes. By using media in which mouse cells cannot grow, because they lack a particular enzyme, but human cells can, the one human chromosome that contains the gene encoding the needed enzyme, will be retained. By using various media, panels of hybrid cell lines can be established. Each cell line in a panel contains either a single human chromosome or a small number of human chromosomes, and a full set of mouse chromosomes, allowing easy mapping of individual genes to specific human chromosomes. (D'Eustachio P. *et al.* (1983) *Science* 220:919-924). Somatic cell hybrids containing only fragments of human chromosomes can also be produced by using human chromosomes with translocations and deletions.

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the C/SKARP-1 nucleotide sequences to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a C/SKARP-1 sequence to its chromosome include *in situ* hybridization (described in Fan, Y. *et al.* (1990) *Proc. Natl. Acad. Sci. USA*, 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. Chromosome spreads can be made using cells whose division has been blocked in metaphase by a chemical such as colcemid that disrupts the mitotic spindle. The chromosomes can be treated briefly with trypsin, and then stained with Giemsa. A pattern of light and dark bands develops on each chromosome, so that the chromosomes can be identified individually. The FISH technique can be used with a DNA sequence as short as 500 or 600 bases. However, clones larger than 1,000 bases have a higher likelihood of

binding to a unique chromosomal location with sufficient signal intensity for simple detection. Preferably 1,000 bases, and more preferably 2,000 bases will suffice to get good results at a reasonable amount of time. For a review of this technique, see Verma *et al.*, Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York 1988).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between a gene and a disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, for example, Egeland, J. *et al.* (1987) *Nature,* 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with the C/SKARP-1 gene, can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes, such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

2. Tissue Typing

The C/SKARP-1 sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the C/SKARP-1 nucleotide sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The C/SKARP-1 nucleotide sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the noncoding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:3 are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from C/SKARP-1 nucleotide sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

3. Use of Partial C/SKARP-1 Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, *e.g.*, hair or skin, or body fluids, *e.g.*, blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, *e.g.*, PCR primers, targeted to specific loci in the human genome, which can

enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (*i.e.* another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions of SEQ ID NO:1 are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the C/SKARP-1 nucleotide sequences or portions thereof, *e.g.*, fragments derived from the noncoding regions of SEQ ID NO:1 having a length of at least 20 bases, preferably at least 30 bases.

The C/SKARP-1 nucleotide sequences described herein can further be used to provide polynucleotide reagents, *e.g.*, labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, *e.g.*, brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such C/SKARP-1 probes can be used to identify tissue by species and/or by organ type.

In a similar fashion, these reagents, *e.g.*, C/SKARP-1 primers or probes can be used to screen tissue culture for contamination (*i.e.* screen for the presence of a mixture of different types of cells in a culture).

C. Predictive Medicine:

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, and monitoring clinical trials are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining C/SKARP-1 protein and/or nucleic acid expression as well as C/SKARP-1 activity, in the context of a biological sample (*e.g.*, blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant or unwanted C/SKARP-1 expression or activity. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with C/SKARP-1 protein, nucleic acid expression or activity. For example, mutations in a C/SKARP-1 gene can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with C/SKARP-1 protein, nucleic acid expression or activity.

Another aspect of the invention pertains to monitoring the influence of agents (*e.g.*, drugs, compounds) on the expression or activity of C/SKARP-1 in clinical trials.

These and other agents are described in further detail in the following sections.

1. Diagnostic Assays

An exemplary method for detecting the presence or absence of C/SKARP-1 protein, polypeptide or nucleic acid in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting C/SKARP-1 protein, polypeptide or nucleic acid (*e.g.*, mRNA, genomic DNA) that encodes C/SKARP-1 protein such that the presence of C/SKARP-1 protein, polypeptide or nucleic acid is detected in the biological sample. In another aspect, the present invention provides a method for detecting the presence of C/SKARP-1 activity in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of C/SKARP-1 activity such that the presence of C/SKARP-1 activity is detected in the biological sample. A preferred agent for detecting C/SKARP-1 mRNA or genomic DNA is a labeled nucleic acid probe capable of hybridizing to C/SKARP-1 mRNA or genomic DNA. The nucleic acid probe can be, for example, a full-length C/SKARP-1 nucleic acid, such as the nucleic acid of SEQ ID NO:1 or 3, or the DNA insert of the plasmid deposited with ATCC as Accession Number _____, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to C/SKARP-1 mRNA or genomic DNA. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting C/SKARP-1 protein is an antibody capable of binding to C/SKARP-1 protein, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (*e.g.*, Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (*i.e.*, physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect C/SKARP-1 mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of C/SKARP-1 mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of C/SKARP-1 protein include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of C/SKARP-1 genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of C/SKARP-1 protein include introducing into a subject a labeled

anti-C/SKARP-1 antibody. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

The present invention also provides diagnostic assays for identifying the presence or absence of a genetic alteration characterized by at least one of (i) aberrant modification or mutation of a gene encoding a C/SKARP-1 protein; (ii) aberrant expression of a gene encoding a C/SKARP-1 protein; (iii) mis-regulation of the gene; and (iii) aberrant post-translational modification of a C/SKARP-1 protein, wherein a wild-type form of the gene encodes a protein with a C/SKARP-1 activity. "Misexpression or aberrant expression", as used herein, refers to a non-wild type pattern of gene expression, at the RNA or protein level. It includes, but is not limited to, expression at non-wild type levels (*e.g.*, over or under expression); a pattern of expression that differs from wild type in terms of the time or stage at which the gene is expressed (*e.g.*, increased or decreased expression (as compared with wild type) at a predetermined developmental period or stage); a pattern of expression that differs from wild type in terms of decreased expression (as compared with wild type) in a predetermined cell type or tissue type; a pattern of expression that differs from wild type in terms of the splicing size, amino acid sequence, post-transitional modification, or biological activity of the expressed polypeptide; a pattern of expression that differs from wild type in terms of the effect of an environmental stimulus or extracellular stimulus on expression of the gene (*e.g.*, a pattern of increased or decreased expression (as compared with wild type) in the presence of an increase or decrease in the strength of the stimulus).

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a serum sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting C/SKARP-1 protein, mRNA, or genomic DNA, such that the presence of C/SKARP-1 protein, mRNA or genomic DNA is detected in the biological sample, and comparing the presence of C/SKARP-1 protein, mRNA or genomic DNA in the control sample with the presence of C/SKARP-1 protein, mRNA or genomic DNA in the test sample.

The invention also encompasses kits for detecting the presence of C/SKARP-1 in a biological sample. For example, the kit can comprise a labeled compound or agent capable of detecting C/SKARP-1 protein or mRNA in a biological sample; means for determining the amount of C/SKARP-1 in the sample; and means for comparing the amount of C/SKARP-1 in the sample with a standard. The compound or agent can be packaged in a

suitable container. The kit can further comprise instructions for using the kit to detect C/SKARP-1 protein or nucleic acid.

2. Prognostic Assays

The diagnostic methods described herein can furthermore be utilized to identify subjects having or at risk of developing a disease or disorder associated with aberrant or unwanted C/SKARP-1 expression or activity. As used herein, the term "aberrant" includes a C/SKARP-1 expression or activity which deviates from the wild type C/SKARP-1 expression or activity. Aberrant expression or activity includes increased or decreased expression or activity, as well as expression or activity which does not follow the wild type developmental pattern of expression or the subcellular pattern of expression. For example, aberrant C/SKARP-1 expression or activity is intended to include the cases in which a mutation in the C/SKARP-1 gene causes the C/SKARP-1 gene to be under-expressed or over-expressed and situations in which such mutations result in a non-functional C/SKARP-1 protein or a protein which does not function in a wild-type fashion, *e.g.*, a protein which does not interact with a C/SKARP-1 ligand or one which interacts with a non-C/SKARP-1 ligand. As used herein, the term "unwanted" includes an unwanted phenomenon involved in a biological response such as proliferation or differentiation. For example, the term unwanted includes a C/SKARP-1 expression or activity which is undesirable in a subject.

The assays described herein, such as the preceding diagnostic assays or the following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with a misregulation in C/SKARP-1 protein activity or nucleic acid expression, such as a cardiovascular disorder. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing a disorder associated with a misregulation in C/SKARP-1 protein activity or nucleic acid expression, such as a cardiovascular disorder. Thus, the present invention provides a method for identifying a disease or disorder associated with aberrant or unwanted C/SKARP-1 expression or activity in which a test sample is obtained from a subject and C/SKARP-1 protein or nucleic acid (*e.g.*, mRNA or genomic DNA) is detected, wherein the presence of C/SKARP-1 protein or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant or unwanted C/SKARP-1 expression or activity. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (*e.g.*, serum), cell sample, or tissue (*e.g.*, cardiac or skeletal muscle tissue).

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant or unwanted C/SKARP-1 expression or activity. For

example, such methods can be used to determine whether a subject can be effectively treated with an agent for a cardiovascular disorder. Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant or unwanted C/SKARP-1 expression or activity in which a test sample is obtained and C/SKARP-1 protein or nucleic acid expression or activity is detected (*e.g.*, wherein the abundance of C/SKARP-1 protein or nucleic acid expression or activity is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant or unwanted C/SKARP-1 expression or activity).

The methods of the invention can also be used to detect genetic alterations in a C/SKARP-1 gene, thereby determining if a subject with the altered gene is at risk for a disorder characterized by misregulation in C/SKARP-1 protein activity or nucleic acid expression, such as a cardiovascular disorder. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic alteration characterized by at least one of an alteration affecting the integrity of a gene encoding a C/SKARP-1-protein, or the mis-expression of the C/SKARP-1 gene. For example, such genetic alterations can be detected by ascertaining the existence of at least one of 1) a deletion of one or more nucleotides from a C/SKARP-1 gene; 2) an addition of one or more nucleotides to a C/SKARP-1 gene; 3) a substitution of one or more nucleotides of a C/SKARP-1 gene, 4) a chromosomal rearrangement of a C/SKARP-1 gene; 5) an alteration in the level of a messenger RNA transcript of a C/SKARP-1 gene, 6) aberrant modification of a C/SKARP-1 gene, such as of the methylation pattern of the genomic DNA, 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of a C/SKARP-1 gene, 8) a non-wild type level of a C/SKARP-1-protein, 9) allelic loss of a C/SKARP-1 gene, and 10) inappropriate post-translational modification of a C/SKARP-1-protein. As described herein, there are a large number of assays known in the art which can be used for detecting alterations in a C/SKARP-1 gene. A preferred biological sample is a tissue or serum sample isolated by conventional means from a subject.

In certain embodiments, detection of the alteration involves the use of a probe/primer in a polymerase chain reaction (PCR) (see, *e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, *e.g.*, Landegran *et al.* (1988) *Science* 241:1077-1080; and Nakazawa *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in the C/SKARP-1-gene (see Abravaya *et al.* (1995) *Nucleic Acids Res* .23:675-682). This method can include the steps of collecting a sample of cells from a subject, isolating nucleic acid (*e.g.*, genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to a C/SKARP-1 gene under conditions such that hybridization and amplification of the C/SKARP-1-gene (if present) occurs, and detecting the presence or absence of an

amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, D.Y. *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi, P.M. *et al.* (1988) *Bio-Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a C/SKARP-1 gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations in C/SKARP-1 can be identified by hybridizing a sample and control nucleic acids, *e.g.*, DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin, M.T. *et al.* (1996) *Human Mutation* 7: 244-255; Kozal, M.J. *et al.* (1996) *Nature Medicine* 2: 753-759). For example, genetic mutations in C/SKARP-1 can be identified in two dimensional arrays containing light-generated DNA probes as described in Cronin, M.T. *et al. supra.* Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the C/SKARP-1 gene and detect mutations by comparing the sequence of the sample C/SKARP-1 with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxam and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc.*

Natl. Acad. Sci. USA 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Biotechniques* 19:448), including sequencing by mass spectrometry (see, *e.g.*, PCT International Publication No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; and Griffin *et al.* (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in the C/SKARP-1 gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers *et al.* (1985) *Science* 230:1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labeled) RNA or DNA containing the wild-type C/SKARP-1 sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton *et al.* (1988) *Proc. Natl Acad Sci USA* 85:4397; Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in C/SKARP-1 cDNAs obtained from samples of cells. For example, the mutY enzyme of E. coli cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a C/SKARP-1 sequence, *e.g.*, a wild-type C/SKARP-1 sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in C/SKARP-1 genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita *et al.* (1989) *Proc Natl. Acad. Sci USA*: 86:2766, see also Cotton (1993) *Mutat. Res.* 285:125-144; and Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control C/SKARP-1 nucleic acids will

be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen *et al.* (1991) *Trends Genet* 7:5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers *et al.* (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl Acad. Sci USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, *e.g.*, in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a C/SKARP-1 gene.

Furthermore, any cell type or tissue in which C/SKARP-1 is expressed may be utilized in the prognostic assays described herein.

3. Monitoring of Effects During Clinical Trials

Monitoring the influence of agents (*e.g.*, drugs) on the expression or activity of a C/SKARP-1 protein (*e.g.*, the modulation signaling pathways associated with cellular growth and differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent determined by a screening assay as described herein to increase C/SKARP-1 gene expression, protein levels, or upregulate C/SKARP-1 activity, can be monitored in clinical trials of subjects exhibiting decreased C/SKARP-1 gene expression, protein levels, or downregulated C/SKARP-1 activity. Alternatively, the effectiveness of an agent determined by a screening assay to decrease C/SKARP-1 gene expression, protein levels, or downregulate C/SKARP-1 activity, can be monitored in clinical trials of subjects exhibiting increased C/SKARP-1 gene expression, protein levels, or upregulated C/SKARP-1 activity. In such clinical trials, the expression or activity of a C/SKARP-1 gene, and preferably, other genes that have been implicated in, for example, a C/SKARP-1-associated disorder can be used as a "read out" or markers of the phenotype of a particular cell.

For example, and not by way of limitation, genes, including C/SKARP-1, that are modulated in cells by treatment with an agent (*e.g.*, compound, drug or small molecule) which modulates C/SKARP-1 activity (*e.g.*, identified in a screening assay as described herein) can be identified. Thus, to study the effect of agents on C/SKARP-1-associated disorders (*e.g.*, disorders characterized by deregulated cellular growth or differentiation), for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of C/SKARP-1 and other genes implicated in the C/SKARP-1-associated disorder, respectively. The levels of gene expression (*e.g.*, a gene expression pattern) can be quantified by northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of C/SKARP-1 or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (*e.g.*, an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) including the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of expression of a C/SKARP-1 protein, mRNA, or genomic DNA in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level of expression or activity of the C/SKARP-1 protein, mRNA, or genomic DNA in the post-administration samples; (v) comparing the level of expression or activity of the C/SKARP-1 protein, mRNA, or genomic DNA in the pre-administration sample with the C/SKARP-1 protein, mRNA, or genomic DNA in the post administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of C/SKARP-1 to higher levels than detected, *i.e.*, to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of C/SKARP-1 to lower levels than detected, *i.e.* to decrease the effectiveness of the agent. According to such an embodiment, C/SKARP-1 expression or activity may be used as an indicator of the effectiveness of an agent, even in the absence of an observable phenotypic response.

4. Electronic Apparatus Readable Media and Arrays

Electronic apparatus readable media comprising C/SKARP-1 sequence information is also provided. As used herein, "C/SKARP-1 sequence information" refers to any nucleotide and/or amino acid sequence information particular to the C/SKARP-1 molecules of the present invention, including but not limited to full-length nucleotide and/or amino acid sequences, partial nucleotide and/or amino acid sequences, polymorphic sequences including single nucleotide polymorphisms (SNPs), epitope sequences, and the like. Moreover, information "related to" said C/SKARP-1 sequence information includes detection of the presence or absence of a sequence (*e.g.*, detection of expression of a sequence, fragment, polymorphism, etc.), determination of the level of a sequence (*e.g.*, detection of a level of expression, for example, a quantitative detection), detection of a reactivity to a sequence (*e.g.*, detection of protein expression and/or levels, for example, using a sequence-specific antibody), and the like. As used herein, "electronic apparatus readable media" refers to any suitable medium for storing, holding or containing data or information that can be read and accessed directly by an electronic apparatus. Such media can include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as compact disc; electronic storage media such as RAM, ROM, EPROM, EEPROM and the like; general hard disks and

hybrids of these categories such as magnetic/optical storage media. The medium is adapted or configured for having recorded thereon C/SKARP-1 sequence information of the present invention.

As used herein, the term "electronic apparatus" is intended to include any suitable computing or processing apparatus or other device configured or adapted for storing data or information. Examples of electronic apparatus suitable for use with the present invention include stand-alone computing apparatus; networks, including a local area network (LAN), a wide area network (WAN) Internet, Intranet, and Extranet; electronic appliances such as a personal digital assistants (PDAs), cellular phone, pager and the like; and local and distributed processing systems.

As used herein, "recorded" refers to a process for storing or encoding information on the electronic apparatus readable medium. Those skilled in the art can readily adopt any of the presently known methods for recording information on known media to generate manufactures comprising the C/SKARP-1 sequence information.

A variety of software programs and formats can be used to store the sequence information on the electronic apparatus readable medium. For example, the sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and MicroSoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like, as well as in other forms. Any number of dataprocessor structuring formats (*e.g.*, text file or database) may be employed in order to obtain or create a medium having recorded thereon the C/SKARP-1 sequence information.

By providing C/SKARP-1 sequence information in readable form, one can routinely access the sequence information for a variety of purposes. For example, one skilled in the art can use the sequence information in readable form to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of the sequences of the invention which match a particular target sequence or target motif.

The present invention therefore provides a medium for holding instructions for performing a method for determining whether a subject has a C/SKARP-1-associated disease or disorder or a pre-disposition to a C/SKARP-1-associated disease or disorder, wherein the method comprises the steps of determining C/SKARP-1 sequence information associated with the subject and based on the C/SKARP-1 sequence information, determining whether the subject has a C/SKARP-1-associated disease or disorder or a pre-disposition to a C/SKARP-1-associated disease or disorder and/or recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention further provides in an electronic system and/or in a network, a method for determining whether a subject has a C/SKARP-1-associated disease or disorder

or a pre-disposition to a disease associated with a C/SKARP-1 wherein the method comprises the steps of determining C/SKARP-1 sequence information associated with the subject, and based on the C/SKARP-1 sequence information, determining whether the subject has a C/SKARP-1-associated disease or disorder or a pre-disposition to a C/SKARP-1-associated disease or disorder, and/or recommending a particular treatment for the disease, disorder or pre-disease condition. The method may further comprise the step of receiving phenotypic information associated with the subject and/or acquiring from a network phenotypic information associated with the subject.

The present invention also provides in a network, a method for determining whether a subject has a C/SKARP-1-associated disease or disorder or a pre-disposition to a C/SKARP-1-associated disease or disorder associated with C/SKARP-1, said method comprising the steps of receiving C/SKARP-1 sequence information from the subject and/or information related thereto, receiving phenotypic information associated with the subject, acquiring information from the network corresponding to C/SKARP-1 and/or a C/SKARP-1-associated disease or disorder, and based on one or more of the phenotypic information, the C/SKARP-1 information (*e.g.*, sequence information and/or information related thereto), and the acquired information, determining whether the subject has a C/SKARP-1-associated disease or disorder or a pre-disposition to a C/SKARP-1-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The present invention also provides a business method for determining whether a subject has a C/SKARP-1-associated disease or disorder or a pre-disposition to a C/SKARP-1-associated disease or disorder, said method comprising the steps of receiving information related to C/SKARP-1 (*e.g.*, sequence information and/or information related thereto), receiving phenotypic information associated with the subject, acquiring information from the network related to C/SKARP-1 and/or related to a C/SKARP-1-associated disease or disorder, and based on one or more of the phenotypic information, the C/SKARP-1 information, and the acquired information, determining whether the subject has a C/SKARP-1-associated disease or disorder or a pre-disposition to a C/SKARP-1-associated disease or disorder. The method may further comprise the step of recommending a particular treatment for the disease, disorder or pre-disease condition.

The invention also includes an array comprising a C/SKARP-1 sequence of the present invention. The array can be used to assay expression of one or more genes in the array. In one embodiment, the array can be used to assay gene expression in a tissue to ascertain tissue specificity of genes in the array. In this manner, up to about 7600 genes can be simultaneously assayed for expression, one of which can be C/SKARP-1. This allows a profile to be developed showing a battery of genes specifically expressed in one or more tissues.

In addition to such qualitative determination, the invention allows the quantitation of gene expression. Thus, not only tissue specificity, but also the level of expression of a battery of genes in the tissue is ascertainable. Thus, genes can be grouped on the basis of their tissue expression *per se* and level of expression in that tissue. This is useful, for example, in ascertaining the relationship of gene expression between or among tissues. Thus, one tissue can be perturbed and the effect on gene expression in a second tissue can be determined. In this context, the effect of one cell type on another cell type in response to a biological stimulus can be determined. Such a determination is useful, for example, to know the effect of cell-cell interaction at the level of gene expression. If an agent is administered therapeutically to treat one cell type but has an undesirable effect on another cell type, the invention provides an assay to determine the molecular basis of the undesirable effect and thus provides the opportunity to co-administer a counteracting agent or otherwise treat the undesired effect. Similarly, even within a single cell type, undesirable biological effects can be determined at the molecular level. Thus, the effects of an agent on expression of other than the target gene can be ascertained and counteracted.

In another embodiment, the array can be used to monitor the time course of expression of one or more genes in the array. This can occur in various biological contexts, as disclosed herein, for example development of a C/SKARP-1-associated disease or disorder, progression of C/SKARP-1-associated disease or disorder, and processes, such a cellular transformation associated with the C/SKARP-1-associated disease or disorder.

The array is also useful for ascertaining the effect of the expression of a gene on the expression of other genes in the same cell or in different cells (*e.g.*, ascertaining the effect of C/SKARP-1 expression on the expression of other genes). This provides, for example, for a selection of alternate molecular targets for therapeutic intervention if the ultimate or downstream target cannot be regulated.

The array is also useful for ascertaining differential expression patterns of one or more genes in normal and abnormal cells. This provides a battery of genes (*e.g.*, including C/SKARP-1) that could serve as a molecular target for diagnosis or therapeutic intervention.

D. Methods of Treatment:

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant or unwanted C/SKARP-1 expression or activity. With regards to both prophylactic and therapeutic methods of treatment, such treatments may be specifically tailored or modified, based on knowledge obtained from the field of pharmacogenomics. "Pharmacogenomics", as used herein, refers to the application of genomics technologies such as gene sequencing, statistical genetics, and gene expression analysis to drugs in clinical development and on the market. More specifically, the term refers the study of how

a patient's genes determine his or her response to a drug (*e.g.*, a patient's "drug response phenotype", or "drug response genotype".) Thus, another aspect of the invention provides methods for tailoring an individual's prophylactic or therapeutic treatment with either the C/SKARP-1 molecules of the present invention or C/SKARP-1 modulators according to that individual's drug response genotype. Pharmacogenomics allows a clinician or physician to target prophylactic or therapeutic treatments to patients who will most benefit from the treatment and to avoid treatment of patients who will experience toxic drug-related side effects.

Treatment is defined as the application or administration of a therapeutic agent to a patient, or application or administration of a therapeutic agent to an isolated tissue or cell line from a patient, who has a disease, a symptom of disease or a predisposition toward a disease, with the purpose to cure, heal, alleviate, relieve, alter, remedy, ameliorate, improve or affect the disease, the symptoms of disease or the predisposition toward disease.

A therapeutic agent includes, but is not limited to, small molecules, peptides, antibodies, ribozymes and antisense oligonucleotides.

1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant or unwanted C/SKARP-1 expression or activity, by administering to the subject a C/SKARP-1 or an agent which modulates C/SKARP-1 expression or at least one C/SKARP-1 activity. Subjects at risk for a disease which is caused or contributed to by aberrant or unwanted C/SKARP-1 expression or activity can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the C/SKARP-1 aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of C/SKARP-1 aberrancy, for example, a C/SKARP-1, C/SKARP-1 agonist or C/SKARP-1 antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

2. Therapeutic Methods

Another aspect of the invention pertains to methods of modulating C/SKARP-1 expression or activity for therapeutic purposes (*e.g.*, for treating subjects having a cardiovascular disease or disorder, for example, congestive heart failure or cardiomyopathy). Accordingly, in an exemplary embodiment, the modulatory method of the invention involves contacting a cell capable of expressing C/SKARP-1 with an agent that modulates one or more of the activities of C/SKARP-1 protein activity associated with the cell, such that C/SKARP-1 activity in the cell is modulated. An agent that modulates C/SKARP-1

protein activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring target molecule of a C/SKARP-1 protein (*e.g.*, a C/SKARP-1 substrate), a C/SKARP-1 antibody, a C/SKARP-1 agonist or antagonist, a peptidomimetic of a C/SKARP-1 agonist or antagonist, or other small molecule. In one embodiment, the agent stimulates one or more C/SKARP-1 activities. Examples of such stimulatory agents include active C/SKARP-1 protein and a nucleic acid molecule encoding C/SKARP-1 that has been introduced into the cell. In another embodiment, the agent inhibits one or more C/SKARP-1 activities. Examples of such inhibitory agents include antisense C/SKARP-1 nucleic acid molecules, anti-C/SKARP-1 antibodies, and C/SKARP-1 inhibitors. These modulatory methods can be performed *in vitro* (*e.g.*, by culturing the cell with the agent) or, alternatively, *in vivo* (*e.g.*, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant or unwanted expression or activity of a C/SKARP-1 protein or nucleic acid molecule. In one embodiment, the method involves administering an agent (*e.g.*, an agent identified by a screening assay described herein), or combination of agents that modulates (*e.g.*, upregulates or downregulates) C/SKARP-1 expression or activity. In another embodiment, the method involves administering a C/SKARP-1 protein or nucleic acid molecule as therapy to compensate for reduced, aberrant, or unwanted C/SKARP-1 expression or activity.

Stimulation of C/SKARP-1 activity is desirable in situations in which C/SKARP-1 is abnormally downregulated and/or in which increased C/SKARP-1 activity is likely to have a beneficial effect. For example, stimulation of C/SKARP-1 activity is desirable in situations in which a C/SKARP-1 is downregulated and/or in which increased C/SKARP-1 activity is likely to have a beneficial effect. Likewise, inhibition of C/SKARP-1 activity is desirable in situations in which C/SKARP-1 is abnormally upregulated and/or in which decreased C/SKARP-1 activity is likely to have a beneficial effect.

3. Pharmacogenomics

The C/SKARP-1 molecules of the present invention, as well as agents, or modulators which have a stimulatory or inhibitory effect on C/SKARP-1 activity (*e.g.*, C/SKARP-1 gene expression) as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) C/SKARP-1-associated disorders (*e.g.*, cardiovascular disorders) associated with aberrant or unwanted C/SKARP-1 activity. In conjunction with such treatment, pharmacogenomics (*i.e.*, the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, a physician or clinician may consider applying

knowledge obtained in relevant pharmacogenomics studies in determining whether to administer a C/SKARP-1 molecule or C/SKARP-1 modulator as well as tailoring the dosage and/or therapeutic regimen of treatment with a C/SKARP-1 molecule or C/SKARP-1 modulator.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. See, for example, Eichelbaum, M. *et al.* (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11) :983-985 and Linder, M.W. *et al.* (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body (altered drug action) or genetic conditions transmitted as single factors altering the way the body acts on drugs (altered drug metabolism). These pharmacogenetic conditions can occur either as rare genetic defects or as naturally-occurring polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

One pharmacogenomics approach to identifying genes that predict drug response, known as "a genome-wide association", relies primarily on a high-resolution map of the human genome consisting of already known gene-related markers (*e.g.*, a "bi-allelic" gene marker map which consists of 60,000-100,000 polymorphic or variable sites on the human genome, each of which has two variants.) Such a high-resolution genetic map can be compared to a map of the genome of each of a statistically significant number of patients taking part in a Phase II/III drug trial to identify markers associated with a particular observed drug response or side effect. Alternatively, such a high resolution map can be generated from a combination of some ten-million known single nucleotide polymorphisms (SNPs) in the human genome. As used herein, a "SNP" is a common alteration that occurs in a single nucleotide base in a stretch of DNA. For example, a SNP may occur once per every 1000 bases of DNA. A SNP may be involved in a disease process, however, the vast majority may not be disease-associated. Given a genetic map based on the occurrence of such SNPs, individuals can be grouped into genetic categories depending on a particular pattern of SNPs in their individual genome. In such a manner, treatment regimens can be tailored to groups of genetically similar individuals, taking into account traits that may be common among such genetically similar individuals.

Alternatively, a method termed the "candidate gene approach", can be utilized to identify genes that predict drug response. According to this method, if a gene that encodes a drugs target is known (*e.g.*, a C/SKARP-1 protein of the present invention), all common variants of that gene can be fairly easily identified in the population and it can be determined

if having one version of the gene versus another is associated with a particular drug response.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (*e.g.*, N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, PM show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification.

Alternatively, a method termed the "gene expression profiling", can be utilized to identify genes that predict drug response. For example, the gene expression of an animal dosed with a drug (*e.g.*, a C/SKARP-1 molecule or C/SKARP-1 modulator of the present invention) can give an indication whether gene pathways related to toxicity have been turned on.

Information generated from more than one of the above pharmacogenomics approaches can be used to determine appropriate dosage and treatment regimens for prophylactic or therapeutic treatment an individual. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a C/SKARP-1 molecule or C/SKARP-1 modulator, such as a modulator identified by one of the exemplary screening assays described herein.

4. Use of C/SKARP-1 Molecules as Surrogate Markers

The C/SKARP-1 molecules of the invention are also useful as markers of disorders or disease states, as markers for precursors of disease states, as markers for predisposition of disease states, as markers of drug activity, or as markers of the pharmacogenomic profile of a subject. Using the methods described herein, the presence, absence and/or quantity of the C/SKARP-1molecules of the invention may be detected, and may be correlated with one or

more biological states *in vivo*. For example, the C/SKARP-1 molecules of the invention may serve as surrogate markers for one or more disorders or disease states or for conditions leading up to disease states. As used herein, a "surrogate marker" is an objective biochemical marker which correlates with the absence or presence of a disease or disorder, or with the progression of a disease or disorder (*e.g.*, with the presence or absence of a tumor). The presence or quantity of such markers is independent of the disease. Therefore, these markers may serve to indicate whether a particular course of treatment is effective in lessening a disease state or disorder. Surrogate markers are of particular use when the presence or extent of a disease state or disorder is difficult to assess through standard methodologies (*e.g.*, early stage tumors), or when an assessment of disease progression is desired before a potentially dangerous clinical endpoint is reached (*e.g.*, an assessment of cardiovascular disease may be made using cholesterol levels as a surrogate marker, and an analysis of HIV infection may be made using HIV RNA levels as a surrogate marker, well in advance of the undesirable clinical outcomes of myocardial infarction or fully-developed AIDS). Examples of the use of surrogate markers in the art include: Koomen *et al.* (2000) *J. Mass. Spectrom.* 35: 258-264; and James (1994) *AIDS Treatment News Archive* 209.

The C/SKARP-1 molecules of the invention are also useful as pharmacodynamic markers. As used herein, a "pharmacodynamic marker" is an objective biochemical marker which correlates specifically with drug effects. The presence or quantity of a pharmacodynamic marker is not related to the disease state or disorder for which the drug is being administered; therefore, the presence or quantity of the marker is indicative of the presence or activity of the drug in a subject. For example, a pharmacodynamic marker may be indicative of the concentration of the drug in a biological tissue, in that the marker is either expressed or transcribed or not expressed or transcribed in that tissue in relationship to the level of the drug. In this fashion, the distribution or uptake of the drug may be monitored by the pharmacodynamic marker. Similarly, the presence or quantity of the pharmacodynamic marker may be related to the presence or quantity of the metabolic product of a drug, such that the presence or quantity of the marker is indicative of the relative breakdown rate of the drug *in vivo*. Pharmacodynamic markers are of particular use in increasing the sensitivity of detection of drug effects, particularly when the drug is administered in low doses. Since even a small amount of a drug may be sufficient to activate multiple rounds of marker (*e.g.*, a C/SKARP-1marker) transcription or expression, the amplified marker may be in a quantity which is more readily detectable than the drug itself. Also, the marker may be more easily detected due to the nature of the marker itself; for example, using the methods described herein, anti- C/SKARP-1antibodies may be employed in an immune-based detection system for a C/SKARP-1 protein marker, or C/SKARP-1-specific radiolabeled probes may be used to detect a C/SKARP-1 mRNA marker. Furthermore, the use of a pharmacodynamic marker may offer mechanism-based

prediction of risk due to drug treatment beyond the range of possible direct observations. Examples of the use of pharmacodynamic markers in the art include: Matsuda *et al.* US 6,033,862; Hattis *et al.* (1991) *Env. Health Perspect.* 90: 229-238; Schentag (1999) *Am. J. Health-Syst. Pharm.* 56 Suppl. 3: S21-S24; and Nicolau (1999) *Am, J. Health-Syst. Pharm.* 56 Suppl. 3: S16-S20.

The C/SKARP-1 molecules of the invention are also useful as pharmacogenomic markers. As used herein, a "pharmacogenomic marker" is an objective biochemical marker which correlates with a specific clinical drug response or susceptibility in a subject (see, *e.g.*, McLeod *et al.* (1999) *Eur. J. Cancer* 35(12): 1650-1652). The presence or quantity of the pharmacogenomic marker is related to the predicted response of the subject to a specific drug or class of drugs prior to administration of the drug. By assessing the presence or quantity of one or more pharmacogenomic markers in a subject, a drug therapy which is most appropriate for the subject, or which is predicted to have a greater degree of success, may be selected. For example, based on the presence or quantity of RNA, or protein (*e.g.*, C/SKARP-1 protein or RNA) for specific tumor markers in a subject, a drug or course of treatment may be selected that is optimized for the treatment of the specific tumor likely to be present in the subject. Similarly, the presence or absence of a specific sequence mutation in C/SKARP-1 DNA may correlate C/SKARP-1 drug response. The use of pharmacogenomic markers therefore permits the application of the most appropriate treatment for each subject without having to administer the therapy.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are incorporated herein by reference.

EXAMPLES

EXAMPLE 1: IDENTIFICATION AND CHARACTERIZATION OF HUMAN C/SKARP-1 cDNA

In this example, the identification and characterization of the gene encoding human C/SKARP-1 (also referred to as clone Fbh33358) is described.

Isolation of the human C/SKARP-1 cDNA

The invention is based, at least in part, on the discovery of genes encoding novel members of the acetyltransferase family.

The nucleotide sequences encoding the human C/SKARP-1 protein is shown in Figure 1 and is set forth as SEQ ID NO:1. The C/SKARP-1 protein encoded by this nucleic acid comprises about 323 amino acids and has the amino acid sequence shown in Figure 1 and set forth as SEQ ID NO:2. The C/SKARP-1 coding region (open reading frame) of SEQ ID NO:1 is set forth as SEQ ID NO:3. Clone Fbh33358 comprising the human C/SKARP-1 cDNA was deposited with the American Type Culture Collection (ATCC®), 10801 University Boulevard, Manassas, VA 20110-2209, on _____, and assigned Accession No. _____.

Analysis of the human C/SKARP-1 Molecules

A search was performed against the HMM database resulting in the identification of six ankyrin repeats (*i.e.*, an ankyrin repeat domain) in the amino acid sequence of human C/SKARP-1 (SEQ ID NO:2) at about residues 64-259 (score: 103.3) of SEQ ID NO:2. Six "ankyrin domains" ("Ank domain") were identified in the amino acid sequence of C/SKARP-1 (SEQ ID NO:2) at about residues 64-96 (score: 17.3); at about residues 97-129 (score: 24.7); at about residues 130-162 (score: 16.4); at about residues 165-194 (score: 12.0); at about residues 195-227 (score: 20.7); and at about residues 229-259 (score: 27.3).

C/SKARP-1 also includes potential casein kinase II phosphorylation sites, for example, from about amino acid residues 101-104, 239-242, 263-266, and 272-275 of SEQ ID NO:2. A potential tyrosine kinase phosphorylation site is found, for example, from about amino acid residues 50-56 of SEQ ID NO:2. Potential N-myristoylation sites are found, for example, from about amino acid residues 58-63, 88-93, 108-113, 121-126, and 142-147 of SEQ ID NO:2. Dileucine motifs are found, for example, from about amino acid residues 26-27, 34-35, 78-79, 117-118, 150-151, 182-183, 215-216, 246-247, 278-279, and 279-280 of SEQ ID NO:2. A potential signal peptide is found, for example, within the first 70 amino acids (amino acid 1 to amino acid 70), of SEQ ID NO:2

Further domain motifs were identified by using the amino acid sequence of C/SKARP-1 (SEQ ID NO:2) to search through the ProDom database. Numerous matches against protein domains described as "ankyrin repeat chromosome XV reading frame", "ankyrin precursor kinase domain signal inhibitor EGF-like", "ankyrin protein cytoskeleton alternative splicing phosphorylation UNC-44 multigene", "F22G12.4 protein", "F34D10.6 protein", "hypothetical 57.7 kD protein", "COL-O putative RNA helicase A", and "mouse BAC library complete BAC-284H12 12P13", and the like were identified.

Tissue Distribution of C/SKARP-1 mRNA

This example describes the tissue distribution of C/SKARP-1 mRNA, as determined by RT-PCR, and as may be determined By Northern blot analysis.

Various cDNA libraries were analyzed by RT-PCR using a human C/SKARP-specific probe. From this analysis it was determined that C/SKARP-1 mRNA was expressed predominantly in heart libraries, from both normal and congestive heart failure samples. C/SKARP-1 mRNA was found to a lesser extent in melanocytes and esophagus (see Figure 2).

Northern blot hybridizations with the various RNA samples would be performed under standard conditions and washed under stringent conditions, *i.e.*, 0.2XSSC at 65°C. The DNA probe was radioactively labeled with ^{32}P-dCTP (using the Prime-It kit (Stratagene, La Jolla, CA) according to the instructions of the supplier). Filters containing human tissue mRNA (MultiTissue Northern I and MultiTissue Northern II from Clontech, Palo Alto, CA) were probed in ExpressHyb hybridization solution (Clontech) and washed at high stringency according to manufacturer's recommendations.

EXAMPLE 2: EXPRESSION OF RECOMBINANT C/SKARP-1 PROTEIN IN BACTERIAL CELLS

In this example, C/SKARP-1 is expressed as a recombinant glutathione-S-transferase (GST) fusion polypeptide in *E. coli* and the fusion polypeptide is isolated and characterized. Specifically, C/SKARP-1 is fused to GST and this fusion polypeptide is expressed in *E. coli*, *e.g.*, strain PEB199. Expression of the GST-C/SKARP-1 fusion protein in PEB199 is induced with IPTG. The recombinant fusion polypeptide is purified from crude bacterial lysates of the induced PEB199 strain by affinity chromatography on glutathione beads. Using polyacrylamide gel electrophoretic analysis of the polypeptide purified from the bacterial lysates, the molecular weight of the resultant fusion polypeptide is determined.

EXAMPLE 3: EXPRESSION OF RECOMBINANT C/SKARP-1 PROTEIN IN COS CELLS

To express the C/SKARP-1 gene in COS cells, the pcDNA/Amp vector by Invitrogen Corporation (San Diego, CA) is used. This vector contains an SV40 origin of replication, an ampicillin resistance gene, an *E. coli* replication origin, a CMV promoter followed by a polylinker region, and an SV40 intron and polyadenylation site. A DNA fragment encoding the entire C/SKARP-1 protein and an HA tag (Wilson *et al.* (1984) *Cell* 37:767) or a FLAG tag fused in-frame to its 3' end of the fragment is cloned into the polylinker region of the vector, thereby placing the expression of the recombinant protein under the control of the CMV promoter.

To construct the plasmid, the C/SKARP-1 DNA sequence is amplified by PCR using two primers. The 5' primer contains the restriction site of interest followed by

approximately twenty nucleotides of the C/SKARP-1 coding sequence starting from the initiation codon; the 3' end sequence contains complementary sequences to the other restriction site of interest, a translation stop codon, the HA tag or FLAG tag and the last 20 nucleotides of the C/SKARP-1 coding sequence. The PCR amplified fragment and the pCDNA/Amp vector are digested with the appropriate restriction enzymes and the vector is dephosphorylated using the CIAP enzyme (New England Biolabs, Beverly, MA). Preferably the two restriction sites chosen are different so that the C/SKARP-1 gene is inserted in the correct orientation. The ligation mixture is transformed into *E. coli* cells (strains HB101, DH5a, SURE, available from Stratagene Cloning Systems, La Jolla, CA, can be used), the transformed culture is plated on ampicillin media plates, and resistant colonies are selected. Plasmid DNA is isolated from transformants and examined by restriction analysis for the presence of the correct fragment.

COS cells are subsequently transfected with the C/SKARP-1-pcDNA/Amp plasmid DNA using the calcium phosphate or calcium chloride co-precipitation methods, DEAE-dextran-mediated transfection, lipofection, or electroporation. Other suitable methods for transfecting host cells can be found in Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory,* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989. The expression of the C/SKARP-1 polypeptide is detected by radiolabelling (^{35}S-methionine or ^{35}S-cysteine available from NEN, Boston, MA, can be used) and immunoprecipitation (Harlow, E. and Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1988) using an HA specific monoclonal antibody. Briefly, the cells are labeled for 8 hours with ^{35}S-methionine (or ^{35}S-cysteine). The culture media are then collected and the cells are lysed using detergents (RIPA buffer, 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM Tris, pH 7.5). Both the cell lysate and the culture media are precipitated with an HA specific monoclonal antibody. Precipitated polypeptides are then analyzed by SDS-PAGE.

Alternatively, DNA containing the C/SKARP-1 coding sequence is cloned directly into the polylinker of the pCDNA/Amp vector using the appropriate restriction sites. The resulting plasmid is transfected into COS cells in the manner described above, and the expression of the C/SKARP-1 polypeptide is detected by radiolabelling and immunoprecipitation using a C/SKARP-1 specific monoclonal antibody.

EXAMPLE 4: TISSUE DISTRIBUTION OF HUMAN C/SKARP-1 mRNA USING TAQMAN™ ANALYSIS

This example describes the tissue distribution of human C/SKARP-1 mRNA in a variety of cells and tissues, as determined using the TaqMan™ procedure. The Taqman™ procedure is a quantitative, reverse transcription PCR-based approach for detecting mRNA.

The RT-PCR reaction exploits the 5' nuclease activity of AmpliTaq Gold™ DNA Polymerase to cleave a TaqMan™ probe during PCR. Briefly, cDNA was generated from the samples of interest, *e.g.*, various human tissue samples, and used as the starting material for PCR amplification. In addition to the 5' and 3' gene-specific primers, a gene-specific oligonucleotide probe (complementary to the region being amplified) was included in the reaction (*i.e.*, the Taqman™ probe). The TaqMan™ probe includes the oligonucleotide with a fluorescent reporter dye covalently linked to the 5' end of the probe (such as FAM (6-carboxyfluorescein), TET (6-carboxy-4,7,2',7'-tetrachlorofluorescein), JOE (6-carboxy-4,5-dichloro-2,7-dimethoxyfluorescein), or VIC) and a quencher dye (TAMRA (6-carboxy-N,N,N',N'-tetramethylrhodamine) at the 3' end of the probe.

During the PCR reaction, cleavage of the probe separates the reporter dye and the quencher dye, resulting in increased fluorescence of the reporter. Accumulation of PCR products is detected directly by monitoring the increase in fluorescence of the reporter dye. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence. During PCR, if the target of interest is present, the probe specifically anneals between the forward and reverse primer sites. The 5'-3' nucleolytic activity of the AmpliTaq™ Gold DNA Polymerase cleaves the probe between the reporter and the quencher only if the probe hybridizes to the target. The probe fragments are then displaced from the target, and polymerization of the strand continues. The 3' end of the probe is blocked to prevent extension of the probe during PCR. This process occurs in every cycle and does not interfere with the exponential accumulation of product. RNA was prepared using the trizol method and treated with DNase to remove contaminating genomic DNA. cDNA was synthesized using standard techniques. Mock cDNA synthesis in the absence of reverse transcriptase resulted in samples with no detectable PCR amplification of the control gene confirms efficient removal of genomic DNA contamination.

Strong expression of C/SKARP-1 mRNA was detected in normal skeletal muscle tissue (set forth in Table 1). In addition, C/SKARP-1 expression was elevated in chronic heart failure tissue as compared with normal heart tissue.

Table 1. Human C/SKARP-1 Taqman Data

Tissue Type	Mean	ß 2 Mean	Ct	Expression
Artery normal	39.62	22.16	17.47	0
Aorta diseased	35.32	22.2	13.12	0
Vein normal	40	20.08	19.92	0
Coronary SMC	40	20.52	19.48	0
HUVEC	38.43	20.94	17.49	0
Hemangioma	36.05	19.48	16.57	0
Heart normal	25.61	20.5	5.11	29.0564
Heart CHF	25.06	20.79	4.27	51.8325
Kidney	38.1	19.56	18.55	0
Skeletal Muscle	25.63	21.57	4.05	60.1622
Adipose normal	39.63	20.51	19.11	0
Pancreas	37	21.78	15.22	0
primary osteoblasts	40	20.2	19.8	0
Osteoclasts (diff)	39.5	17.3	22.2	0
Skin normal	35.94	22.04	13.9	0
Spinal cord normal	39.89	20.77	19.12	0
Brain Cortex normal	38.08	21.66	16.41	0
Brain Hypothalamus normal	39.67	22.25	17.42	0
Nerve	29.73	21.63	8.11	3.6195
DRG (Dorsal Root Ganglion)	38.45	21.3	17.16	0
Breast normal	39.72	20.61	19.1	0
Breast tumor	37.32	20.47	16.85	0
Ovary normal	39.22	19.45	19.77	0
Ovary Tumor	38.84	18.43	20.41	0
Prostate Normal	39.03	19.21	19.82	0
Prostate Tumor	39.73	19.95	19.79	0
Salivary glands	38.84	19.24	19.6	0
Colon normal	37.79	18.52	19.27	0
Colon Tumor	32.63	21.16	11.48	0.3513
Lung normal	35.8	18.08	17.72	0
Lung tumor	29.53	20.31	9.22	1.6769
Lung COPD	36.16	18.25	17.91	0
Colon IBD	37.62	17.41	20.21	0
Liver normal	39.89	19.82	20.07	0
Liver fibrosis	38.04	20.44	17.6	0
Spleen normal	37.47	18.27	19.2	0
Tonsil normal	35.08	18.27	16.81	0
Lymph node normal	38.41	19.79	18.63	0
Small intestine normal	39.95	19.66	20.3	0
Macrophages	40	16.84	23.16	0
Synovium	40	19.55	20.45	0
BM-MNC	40	18.5	21.5	0
Activated PBMC	39.78	17.59	22.18	0
Neutrophils	40	17.78	22.22	0
Megakaryocytes	40	18.25	21.75	0
Erythroid	38.59	20.45	18.15	0
positive control	30.01	20.18	9.84	1.0949

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.